

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7:	
C07K 14/705, C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N 33/53, A61K 38/17	

(11) International Publication Number:

WO 00/40614

(43) International Publication Date:

13 July 2000 (13.07.00)

(21) International Application Number:

PCT/US99/29996

A₂

(22) International Filing Date:

20 December 1999 (20.12.99)

CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE,

(30) Priority Data:

60/114,220 60/120,018 60/140,415 30 December 1998 (30.12.98) US 29 January 1999 (29.01.99) US 22 June 1999 (22.06.99) US Published

Without international search report and to be republished upon receipt of that report.

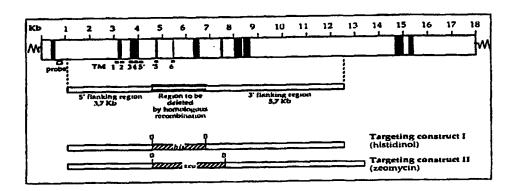
(71) Applicant (for all designated States except US): BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 330 Brookline Avenue, Boston, MA 02215 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): SCHARENBERG, Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA 02420 (US).

(74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(54) Title: CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY



(57) Abstract

Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
ВВ	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	LE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	ſL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	LS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Кепуа	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD.	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

Field of the Invention

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

5

10

15

20

25

30

Background of the Invention

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca²⁺ ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca²⁺ ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca²⁺ cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed store operated calcium influx, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

-2-Summary of the Invention

5

10

15

20

25

30

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca2+ flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca²⁺ from intracellular calcium stores, allowing Ca²⁺ influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25

30

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

WO 00/40614 PCT/US99/29996 -4-

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

- a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;

5

10

15

20

25

30

- c) isolating the SOC/CRAC molecule from the complex; and
- d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

WO 00/40614 PCT/US99/29996 -5-

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

5

10

15

20

25

30

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEO. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)2, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules in vivo or in vitro.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;
- b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

WO 00/40614 PCT/US99/29996
-6-

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

5

10

15

20

25

30

- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

- a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and
- b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-3/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

5

10

15

20

25

30

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, *inter alia*, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEO ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

5

10

15

20

25

30

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a C. Elegans polypeptide at the c05c12.3 locus.

SEQ ID NO:14 is the amino acid sequence of a *C. Elegans* polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a C. Elegans polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

5

10

15

20

25

30

WO 00/40614 PCT/US99/29996 -9-

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEO ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEO ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEO ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEQ ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

Figure 1 is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca2+ flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca²⁺ from intracellular calcium stores, allowing Ca²⁺ influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

5

10

15

20

25

30

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca²⁺ fluxing and, in particular, lymphocyte Ca²⁺ fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

5

10

15

20

25

30

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid" refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca²⁺ transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

-12-

Science, 1992; 256:498-502; and Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

- (a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25

30

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

5

10

15

20

25

30

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified in vitro by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

-14-

5

10

15

20

25

30

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEO ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEQ ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, The term "stringent conditions" as used herein refers to under stringent conditions. parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

-15-

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

5

10

15

20

25

30

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (ftp:/ncbi.nlm.nih.gov/pub/). Exemplary tools include the BLAST system available at http://www.ncbi.nlm.nih.gov. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVetor sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphoimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC cDNA clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

-16-

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

5

10

15

20

25

30

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

-17-

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

5

10

15

20

25

30

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEO ID NO:29, and SEO ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique -18-

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

5

10

15

20

25

30

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., Nat. Med. 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5,SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue experimentation.

5

10

15

20

25

30

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred internucleoside linkages phosphorothioates, alkylphosphonates, synthetic are phosphoramidates, phosphorodithioates, phosphate esters, alkylphosphonothioates, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

5

10

15

20

25

30

-20-

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-Oalkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as <u>E.coli</u> and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

-21-

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

5

10

15

20

25

30

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed -22-

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

5

10

15

20

25

30

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the group consisting of SEO ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

WO 00/40614 PCT/US99/29996 -23-

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

5

10

15

20

25

30

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1α, which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., <u>E. coli</u>), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

-24-

"knock-outs" in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

5

10

15

20

25

30

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca²⁺ fluxing, high selectivity, a unitary

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

5

10

15

20

25

30

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in Science 278:82-87, 1997, whereby proteins can be designed de novo. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

-26-

Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

5

10

15

20

25

30

Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., <u>E. coli</u>, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

-27-

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

5

10

15

20

25

30

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described First, the invention permits isolation of SOC/CRAC polypeptides, elsewhere herein. including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC The polypeptide may be purified from cells which naturally produce the molecules. polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

-28-

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

5

10

15

20

25

30

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

5

10

15

20

25

30

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

5

. 10

15

20

25

30

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

WO 00/40614 PCT/US99/29996 -31-

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

5

10

15

20

25

30

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

WO 00/40614 PCT/US99/29996 -32-

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

5

10

15

20

25

30

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention. a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

5

10

15

20

25

30

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

-34-

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. Trans-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such trans-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic nonhuman animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

5

10

15

20

25

30

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca²⁺ fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

5

10

15

20

25

30

درک

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca²⁺. Such separation is preferred so that a change in Ca²⁺ concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cell-membrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in <u>Capacitative Calcium Entry</u>, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel activation and a given reference calcium concentration. Detection of a decrease in the

WO 00/40614 PCT/US99/29996
-36-

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5

10

15

20

25

30

SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or



polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

5

10

15

20

25

30

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include agents such as dihydropyridines (DHPs), phenylalkylamines, omega conotoxin (omega.-CgTx) and pyrazonoylguanidines.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

-38-

5

10

15

20

25

30

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromotograpic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β-galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

5

10

15

20

25

30

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and Dictyostelium myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., Proc Natl Acad Sci U S A, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or Dictyostelium MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

-40-

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10⁷ M⁻¹, more preferably at least about 10⁸ M⁻¹, and most preferably at least about 10⁹ M⁻¹. The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

5

10

15

20

25

30

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced in vitro or in vivo in a host. Such techniques include transfection of nucleic acid-CaPO₄ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

-41-

polyesteramides, polyorthoesters, polyhydroxybutyric polyanhydrides. acid, and Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sylastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

5

10

15

20

25

30

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, are used herein, means that the implant is constructed and arranged to delivery therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing ex vivo gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction in vitro of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. In vivo gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

-42-

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

5

10

15

20

25

30

As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode C. elegans is nearing completion, and IP3-dependent calcium signals have recently been shown to be required for one or more aspects of C. elegans development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily excluding all proteins below approximately 1000 amino acids in size, followed by focusing on remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated CO5C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was noteable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

We then used C05C12.3 for BLAST alignment screening of the rest of the *C. elegans* genome and also mammalian databases for homologous proteins, revealing two other *C. elegans* homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and

-43-

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRAClike currents were best defined. Subsequent BLAST searches picked up mouse EST sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures

5

10

15

20

25

30

Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

-44-

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promers were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

5

10

15

20

25

30

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y., N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in <u>Capacitative Calcium Entry</u>, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2+ indicators with greatly improved fluorescence properties, J. Biol

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

5

10

15

20

25

30

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2⁺ indicators with greatly improved fluorescence properties, <u>J Biol Chem 260</u>, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

-46-

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., *EMBO J*, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (*Cell*, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure 1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

5

10

15

20

25

30

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

WO 00/40614 PCT/US99/29996
-47-

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in Xenopus oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: Pharmacologic Behavior of SOC/CRAC

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: Gene targeting

5

10

15

20

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectivelly refered to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407,

AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532,

AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170,

AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID

NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333,

AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

-48-Claims

1. An isolated nucleic acid molecule, comprising:

5

10

15

20

25

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).
- 2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.
- 3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.
- 4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.
- 5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.
- 6. An isolated nucleic acid molecule selected from the group consisting of
- (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,
- (b) complements of (a),
 provided that the unique fragment includes a sequence of contiguous nucleotides which is not
 identical to any sequence selected from a sequence group consisting of
 - (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,
 - (2) complements of (1), and
 - (3) fragments of (1) and (2).

_49-

WO 00/40614

5

7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

PCT/US99/29996

- (1) at least two contiguous nucleotides nonidentical to the sequence group,
- (2) at least three contiguous nucleotides nonidentical to the sequence group,
- (3) at least four contiguous nucleotides nonidentical to the sequence group,
- (4) at least five contiguous nucleotides nonidentical to the sequence group,
- (5) at least six contiguous nucleotides nonidentical to the sequence group,
- (6) at least seven contiguous nucleotides nonidentical to the sequence group.
- 8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.
- 9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.
 - 10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.
 - 11. A host cell transformed or transfected with the expression vector of claim 10.
- 20 12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.
 - 13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.
- 14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

- WO 00/40614 PCT/US99/29996 -50-
- 15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.
- 16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.
- 17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:36, SEQ ID NO:30, and SEQ ID NO:32.
 - 18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.
 - 19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.
 - 20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:
 - a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;

5

10

15

20

25

- c) isolating the SOC/CRAC molecule from the complex; and
- d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.
- 21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.
- 22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

- WO 00/40614 PCT/US99/29996
- 23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.
- 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

5

10

15

20

25

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;
- b) detecting a Ca²⁺ concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.
- 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:
- a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and
- b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.
- 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.
- 27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.
- 28. The method of claim 25, wherein the test sample is tissue.
- 29. The method of claim 25, wherein the test sample is a biological fluid.

WO 00/40614 PCT/US99/29996
-52-

30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).

- 31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to SOC/CRAC and polyclonal antisera to SOC/CRAC.
- 32. A kit, comprising a package containing:

5

10

15

20

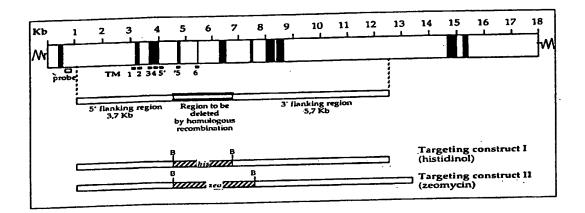
25

30

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

- a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.
- 33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.
- 34. A pharmaceutical composition comprising:
- a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and
 - a pharmaceutically acceptable carrier.
- 35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.
- 36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:
- a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;
- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC kinase activity.
- 37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

FIGURE 1.



-1-SEQUENCE LISTING

```
<110> Beth Israel Deaconess Medical Center, Inc.
            Scharenberg, Andrew
      <120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY
      <130> B0662/7026WO/ERP/KA
      <150> U.S. 60/114,220
      <151> 1998-12-30
      <150> U.S. 60/120,018
      <151> 1999-01-29
      <150> U.S. 60/140,415
      <151> 1999-06-22
      <160> 32
      <170> FastSEQ for Windows Version 3.0
      <210> 1
      <211> 1212
      <212> DNA
      <213> Homo Sapiens
      <400> 1
gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaag ccacttggaa
                                                                        60
actggaacca aagatcaaga aactgtttgc tctaaagcta cagaaggaga taatacagaa
                                                                       120
tttggagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca
                                                                       180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt
                                                                       240
tetettgetg gatttaetga etgteacaga acttecatte etgtteatte aaaacaagaa
                                                                       300
aaaatcagta gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta
                                                                       360
ataccggttt gtagatttca actaaacaga tatatattat taaatacatt aaactttttt
                                                                       420
agataagatc tacaaagtgg tgatatttgg gactatatca aaaattcaaa aaaatttttc
                                                                       480
ttaagaaaac tgactttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc
                                                                       540
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg
                                                                       600
aagctaatta tatgaactca tttccagcta actgcgataa tgattgattt tgcaaattcc
                                                                       660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt
                                                                       720
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt
                                                                       780
acattcatat atatctgttc tggcatagag agattatgtt gttaaaatca tgttattgaa
                                                                       840
aaaagttatt tcagtgggga aagaggttag ttaacaaaga gattcacagt aacaaatcct
                                                                       900
cctttctgga gggactcttc ctgaccctga gctgcacaac tttgcaacaa attaaagcct
                                                                       960
aaccgaagat gacctcacaa tggcaattta gaactcatgg gagtcaactt acataaacgg
                                                                      1020
tatttgattt ctgataagat agtggaatta ttggttatag atgacaaaat aagtatgttt
                                                                      1080
aaagtgatga tggacataaa aaagttttaa atataaaaca tgagaaaaga aggagatact
                                                                      1140
                                                                      1200
attcaaaaag actggcaaat ttgaaaaaact agaaataaaa aaaaaaaaa aaaatgagcg
gccgcaagct tt
                                                                      1212
      <210> 2
      <211> 141
      <212> PRT
      <213> Homo Sapiens
      <400> 2
Ala Arg Gly Lys Phe Phe Val Ser Thr Pro Ser Gln Pro Ser Cys Lys
                5
                                10
Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys
            20
                                25
```

```
-2-
                      . .
Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe Val Gly His Arg
                            40
Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Lys
    50
                        55
                                             60
Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu Lys Arg Val Ser
                    70
                                         75
Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser Ile Pro Val His
Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro Ser Thr Glu Asp Thr His
                                105
                                                    110
Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Val Cys Arg Phe Gln Leu
        115
                            120
                                                 125
Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn Phe Phe Arg
                        135
      <210> 3
<211> 739
      <212> DNA
      <213> Homo Sapiens
      <220>
      <221> unsure
      <222> (5)...(5)
      <223> UNKNOWN
      <221> unsure
      <222> (21)...(22)
<223> UNKNOWN
      <221> unsure
      <222> (29)...(29)
      <223> UNKNOWN
                                                                         60
tegantaggg gtettecace nneatactng gatgatggtt ggtgaagtet atgeatacga
aattgatgtg tgtgcaaacg attctgttat ccctcaaatc tgtggtcctg ggacgtggtt
                                                                        120
gactccattt cttcaagcag tctacctctt tgwacagtat atcattatgg ttaatcttct
                                                                        180
tattgcattt ytcaacaatg tgtatttaca agtgaaggca atttccaata ttgyatggaa
                                                                        240
gtaccagcgt tatcatttta ttatggctta tcatgagaaa ccagttctgc ctcctccact
                                                                        300
tatcattett agecatatag tttetetgtt ttgetgeata tgtaagagaa gaaagaaaga
                                                                        360
taagacttcc gatggaccaa aacttttctt aacagaagaa gatcaaaaga aacttcatga
                                                                        420
ttttgaagag cagtgtgttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg
                                                                        480
                                                                        540
gagtgaagag agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa
agaagttgga gatccgtgtc aactacataa aaagatcatt acaatcatta gattctcaaa
                                                                        600
ttggccattt gcaagatctt tcagccctga cggtagatac attaaaaaca ctcactggcc
                                                                        660
aaaagcgtcg gaagctagca aagttcataa tgaaatcaca cgagaactga gcatttccaa
                                                                        720
acacttggct caaaacctt
                                                                        739
      <210> 4
      <211> 235
      <212> PRT
      <213> Homo Sapiens
      <221> UNSURE
      <222> (41)...(41)
      <223> UNKNOWN
     <221> UNSURE
     <222> (54)...(54)
```

<223> UNKNOWN

<221> UNSURE <222> (68)...(68) <223> UNKNOWN

<400> 4 Met Met Val Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys Ala Asn 5 10 Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu Thr Pro 25 20 Phe Leu Gln Ala Val Tyr Leu Phe Xaa Gln Tyr Ile Ile Met Val Asn 45 35 40 Leu Leu Ile Ala Phe Xaa Asn Asn Val Tyr Leu Gln Val Lys Ala Ile 55 60 Ser Asn Ile Xaa Trp Lys Tyr Gln Arg Tyr His Phe Ile Met Ala Tyr 75 70 His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile 85 90 Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp Lys Thr 110 100 105 Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu 120 125 115 His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu Lys Asp 140 130 135 Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr Phe Glu 155 150 Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp Pro Cys 170 165 Gln Leu His Lys Lys Ile Ile Thr Ile Ile Arg Phe Ser Asn Trp Pro 180 185 Phe Ala Arg Ser Phe Ser Pro Asp Gly Arg Tyr Ile Lys Asn Thr His 200 205 195 Trp Pro Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg 215 Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn 230

<210> 5

<211> 1579

<212> DNA

<213> Homo Sapiens

<220>

<221> unsure

<222> (368)...(368)

<223> g or c

<221> unsure

<222> (372)...(372)

<223> g or c

<221> unsure

<222> (374)...(374)

<223> g or a

<221> unsure

<222> (375)...(375)

<223> g or c

```
<221> unsure
      <222> (387) ... (387)
      <221> unsure
      <222> (482)...(482)
     <400> 5
acqtcqcctq caqqtaccqq tccggaattc ccgggtcgac ccacgcgtcc ggcatggtgt
                                                                        60
                                                                       120
tqtaaataca cttagctcct ctcttcctca aggtgatctt gaaagtaata atccttttca
tigtaatatt ttaatgaaag atgacaaaga tccccagtgt aatatatttg gtcaagactt
                                                                       180
                                                                       240
acctgcagta ccccagagaa aagaatttaa ttttccagag gctggttcct cttctggtgc
cttattccca agtgctgttt cccctccaga actgcgacag agactacatg gggtagaact
                                                                       300
cttaaaaata tttaataaaa atcaaaaatt aggcagttca tctactagca taccacatct
                                                                       360
gtcatccsca csarscaaat tttttgntag tacaccatct cagccaagtt gcaaaagcca
                                                                       420
cttggaaact ggaaccaaag atcaagaaac tgtttgctct aaagctacag aaggagataa
                                                                       480
                                                                       540
tncagaattt ggagcatttg taggacacag agatagcatg gatttacaga ggtttaaaga
aacatcaaac aagataaaaa tactatccaa taacaatact tctgaaaaca ctttgaaacg
                                                                       600
agtgagttct cttgctggat ttactgactg tcacagaact tccattcctg ttcattcaaa
                                                                       660
acaagaaaaa atcagtagaa ggccatctac cgaagacact catgaagtag attccaaagc
                                                                       720
agctttaata ccggtttgta gatttcaact aaacagatat atattattaa atacattaaa
                                                                       780
cttttttaga taagatctac aaagtggtga tatttgggac tatatcaaaa attcaaaaaa
                                                                       840
atttttctta agaaaactga ctttagcata gtagcagtta cagaaaagtt tcttacagtg
                                                                       900
aatagtcagg aattttaaag aaaaatttat gcagaataaa ggcaggaatc tctttttgtt
                                                                       960
tgaattgaag ctaattatat gaactcattt ccagctaact gcgataatga ttgattttgc
                                                                      1020
                                                                      1080
aaattccctt taaaagcaca cactgacaag acaaaaagct caggaaaaagg cagaaaaatt
actcetttat aatcaaqtat tatatataag tcagtgetca taattttget caagaaaata
                                                                      1140
ttgacttaca ttcatatata tctgttctgg catagagaga ttatgttgtt aaaatcatgt
                                                                      1200
tattgaaaaa agttatttca gtggggaaag aggttagtta acaaagagat tcacagtaac
                                                                      1260
                                                                      1320
aaatoctcot ttotggaggg actottoctg accotgaget gcacaacttt gcaacaaatt
aaaqcctaac cqaaqatqac ctcacaatgg caatttagaa ctcatgggag tcaacttaca
                                                                      1380
taaacggtat ttgatttctg ataagatagt ggaattattg gttatagatg acaaaataag
                                                                      1440
                                                                      1500
tatgtttaaa gtgatgatgg acataaaaaa gttttaaata taaaacatga gaaaagaagg
                                                                      1560
agatactatt caaaaagact ggcaaatttg aaaaactaga aataaaaaaa aaaaaaaaa
                                                                      1579
atgagcggcc gcaagcttt
      <210> 6
      <211> 243
      <212> PRT
      <213> Homo Sapiens
      <220>
      <221> UNSURE
      <222> (103)...(105)
      <223> UNKNOWN
      <221> UNSURE
      <222> (109)...(109)
      <223> UNKNOWN
      <221> UNSURE
      <222> (141)...(141)
      <223> UNKNOWN
      <400> 6
Val Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn
                                    10
                 5
Asn Pro Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln
                              25
          20
Cys Asn Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu
                            40
```

-5-Phe Asn Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser 55 60 Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu 75 70 Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser 85 90 Ile Pro His Leu Ser Ser Xaa Xaa Xaa Lys Phe Phe Xaa Ser Thr Pro 110 100 105 Ser Gln Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln 120 125 115 Glu Thr Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Xaa Glu Phe Gly 135 140 Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu 145 150 155 Thr Ser Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn 165 170 175 Thr Leu Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg 190 180 185 Thr Ser Ile Pro Val His Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro 200 205 195 Ser Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro 220 . 210 215 Val Cys Arg Phe Gln Leu Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn 225 235 230

<210> 7 <211> 3532 <212> DNA <213> Mus Musculus <220> <221> unsure <222> (2420)...(2420) <223> unknown <221> unsure <222> (2434)...(2434) <223> unknown <221> unsure <222> (2461)...(2461) <223> unknown <221> unsure <222> (2466)...(2466) <223> unknown

<221> unsure <222> (2470)...(2470) <223> unknown

<400> 7

Phe Phe Arg

attatggett atcatgaaaa accagteetg eeteeteete ttateateet eageeatata 60 gttteaetgt tttgetgtgt atgeaaaaga agaaagaaag ataagaette egatgggeea 120 aaacttttet taacagaaga agateaaaag aaacteeatg attttgaaga geagtgtgtt 180 gagatgtaet ttgatgagaa agatgaeaaa tteaatteetg ggagtgaaga gagaateegg 240 gteaettttg aaagagtgga geagatgage atteagatta aagaagttgg agategtgte 300 aactacataa aaagateatt acagteetta gatteteaaa ttggteatet geaagateete 360

-6tcagccctaa cagtagatac attgaaaaca cttacagccc agaaagcttc agaagctagt 420 aaaqtqcaca atqaqatcac acqaqaattq agtatttcca aacacttggc tcaqaatctt 480 540 attgatgatg ttcctgtaag acctttgtgg gaagaaccta gtgctgtaaa cacactgagt tcctctctc ctcaaggtga tcgggaaagt aataatcctt ttctttgtaa tatttttatg 600 aaaqatqaaa aaqacccca atataatctq tttqqacaaq atttqcccqt qataccccaq 660 agaaaagaat tcaacattcc agaggctggt tcctcctgtg gtgccttatt cccaagtgct 720 gtttctcccc cagaattacg acagagacga catggggtag aaatgttaaa aatatttaat 780 aaaaatcaaa aattaggcag ttcacctaat agttcaccac atatgtcctc cccaccaacc 840 aaattttctg tgagtacccc atcccagcca agttgcaaaa gtcacttgga atccacaacc 900 aaagatcaag aacccatttt ctataaagct gcagaagggg ataacataga atttggagca 960 tttgtgggac acagagatag tatggactta cagaggttta aagaaacatc aaacaaaata 1020 agagaactqt tatctaatqa tactcctgaa aacactctga aacatqtggg tqctqctgga 1080 tatagtgaat gttgtaagac ttctacttct cttcactcgg tgcaagcaga aagctgtagt 1140 agaaqaqcgt cgacggaaqa ctctccagaa gtcgattcta aagcagcttt gttaccggat 1200 1260 tggttacgag atagaccatc aaacagagaa atgccatctg aaggaggaac attaaatggt cttgcttctc catttaagcc cgttttggat acaaattact attattcagc tgtggaaaga 1320 aataacctga tgaggttgtc acagagtatt cccttcgttc ctgtacctcc acgaggcgag 1380 cctqtcacaq tqtaccqtct qqaqqaqaqt tctcccaqta tactqaataa caqcatqtct 1440 1500 tcatggtctc agctaggcct ctgtgccaaa attgagtttt taagtaaaga ggaaatggaa qqtqqtttac gaagagcagt caaaqtqctq tgtacctggt cagagcacga tatcctgaag 1560 tragggrate tetatateat taagteattt citeetgagg tgataaacae atggteaage 1620 atttataaag aagatacggt tctacatctc tgtctcagag aaatacaaca acagagagca 1680 gcacaaaagc tcacatttgc ctttaatcag atgaaaccca aatccatacc atattctcca 1740 aggttccttg aagttttcct gttgtactgc cattcagcag ggcagtggtt tgctgtagaa 1800 gagtgcatga ctggtgaatt tagaaaatac aacaacaata atggtgatga aatcatteet 1860 acaaatactc tagaagagat catgctagcc tttagccact ggacctatga atataccaga 1920 1980 ggggagttac tggtacttga cttacaagga gtgggagaaa acttgactga cccatctgta ataaaagctg aagaaaaaag atcctgtgac atggtttttg gccctgccaa tctaggagaa 2040 gatgcaataa aaaacttcaa gagccaaaca tccactgtaa ttcttgctgt cgaaagctta 2100 aacttcccag atttgaagag gaatgactac acgcccttga taaaattata tttcctcagg 2160 2220 atgagtcatc agatttgaat cttcaatctg gaaattccac caaagaatca gaagcaacaa attetgtteg tetgatgtta tagtgetgag teattggttt ttgeetaeae tteacaaaag tgtaaetgte agtttteett tegggggaat tgatgatata ggaagatgtg tgcaaaatga 2280 2340 gcttgctggc cccacacata gtctagaggt aatgttctca ttgaaaaacg cctggaggtg 2400 2460 gaggetgeag atgecagtgn aaagtgetag etgneagaga gteagtgete tegggetggt naaggneggn accettgetg etgagagtgg tggttetett cacetggtge aggaceatta accaaagtea agtetteaga tttgattgge tgeteagtea eageceatte agetaaggaa 2520 2580 actaaattgc gcagcttttt aaatggctga agtcttcctc agtttgtgct ctatgataat 2640 gatgttaget eteaactagg tgtttgtgge caegggagaa etaeteetta caattttget 2700 tcacaggcat gttacaaagc ctgcactgaa aaccgtttgt cttccctctc tccctccctc 2760 2820 ttttccctgt agtattgagg atcaaaccca gggcctcatg aagaccattt tctaagagac attttattta aqaatcaact atagagtcta tgtttatgga tacagccagt ttttgttaaa 2880 2940 caaaacctga attgtgcaaa agggtttttt aacatttatc aatgttaagt aaaagaaagc catgataaat aagaattaac tcactgttca atgggtgttt cctgtgagga aggttacagt 3000 tgtaacagcc tgcagttgca tacatctcca aagatttaca gacttagtgt atcaaatcag 3060 agtgtcatgt gagctctcac attgaaaatt ctataggaat gtgtcaatgt gaattctatt 3120 tctggtactt aagaaatcag ttgttggatt atccttatac agtataggga gatcacaata 3180 3240 caactttatg ccaataaaat ctaacttaat tgcccagata tttttgcata tttagcaaca agaaaagctt atcatttgac tcaagtttta tgctttctct ttctttcat ttcctaggta 3300 ctaattttaa tttttatttg gaaggagcag tgtaaagctt acttgtattc aatagtgtat 3360 3420 ctcatagata cagacaaggc cgcagagata agctgttaaa tagtgtttaa tgttgatgtg 3480 gagagaaagg tgtattactt aaaaatacta taccatatac gttttgtata tcattaaatc

3532

<210> 8

<211> 475

<212> PRT

<213> Mus Musculus

<400> 8

Ile Met Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile Val Ser Leu Phe Cys Cys Val Cys Lys Arg Arg Lys Lys Asp Lys Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asp Glu Lys Asp Asp Lys Phe Asn Ser Gly Ser Glu Glu Arg Ile Arg 65 70 75 80 Val Thr Phe Glu Arg Val Glu Gln Met Ser Ile Gln Ile Lys Glu Val 85 90 95 Gly Asp Arg Val Asn Tyr Ile Lys Arg Ser Leu Gln Ser Leu Asp Ser Gln Ile Gly His Leu Gln Asp Leu Ser Ala Leu Thr Val Asp Thr Leu Lys Thr Leu Thr Ala Gln Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn Leu Ile Asp Asp Val Pro Val Arg Pro Leu Trp Glu Glu Pro Ser Ala Val 170 175 Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Arg Glu Ser Asn Asn Pro Phe Leu Cys Asn Ile Phe Met Lys Asp Glu Lys Asp Pro Gln Tyr Asn Leu Phe Gly Gln Asp Leu Pro Val Ile Pro Gln Arg Lys Glu Phe Asn Ile Pro Glu Ala Gly Ser Ser Cys Gly Ala Leu Phe Pro Ser Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Arg His Gly Val Glu Met Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Pro Asn Ser Ser Pro His Met Ser Ser Pro Pro Thr Lys Phe Ser Val Ser Thr Pro Ser Gln Pro Ser Cys Lys Ser His Leu Glu Ser Thr Thr Lys Asp Gln Glu Pro Ile Phe Tyr Lys Ala Ala Glu Gly Asp Asn Ile Glu Phe Gly Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Arg Glu Leu Leu Ser Asn Asp Thr Pro Glu Asn Thr Leu Lys His Val Gly Ala Ala Gly Tyr Ser Glu Cys Cys Lys Thr Ser Thr Ser Leu His Ser Val Gln Ala Glu Ser Cys Ser Arg Arg Ala Ser Thr Glu Asp Ser Pro Glu Val Asp Ser Lys Ala Ala Leu Leu Pro Asp Trp Leu Arg Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Gly Gly Thr Leu Asn Gly Leu Ala Ser Pro Phe Lys Pro Val Leu Asp Thr Asn 425 430 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln Ser Ile Pro Phe Val Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val 455 460 Tyr Pro Ser Gly Gly Arg Val Leu Pro Val Tyr

```
<210> 9
<211> 5433
<212> DNA
<213> Mus Musculus
<220>
<221> unsure
<222> (5094)...(5094)
<223> unknown
```

<400> 9

qqctqaaaqa qcctqaqctq tqcctctcca ttccactqct gtgqcagggt cagaaatctt 60 ggatagagaa aaccttttgc aaacgggaat gtatctttgt aattcctagc acgaaagact 120 ctaacaggtg ttgctgtggc cagttcacca accagcatat cccccctctg ccaagtgcaa 180 240 cacccagcaa aaatgaagag gaaagcaaac aggtggagac tcagcctgag aaatggtctg ttgccaagca cacccagagc tacccaacag attcctatgg agttcttgaa ttccagggtg 300 360 geggatatte caataaagee atgtatatee gtgtateeta tgacaccaag ceagacteae tgctccatct catggtgaaa gattggcagc tggaactccc caagctctta atatctgtgc 420 atggaggeet ccagaacttt gagatgeage ccaagetgaa acaagtettt gggaaaggee 480 tgatcaaggc tgctatgacc accggggcct ggatcttcac cgggggtgtc agcacaggtg 540 ttatcagcca cgtaggggat gccttgaaag accactcctc caagtccaga ggccgggttt 600 gtgctatagg aattgctcca tggggcatcg tggagaataa ggaagacctg gttggaaagg 660 atgtaacaag agtgtaccag accatgtcca accetctaag taagetetet gtgeteaaca 720 780 actoccacac ccacttcatc ctggctgaca atggcaccct gggcaagtat ggcgccgagg tgaagetgeg aaggetgetg gaaaageaca teteceteca gaagateaae acaagaetgg 840 900 ggcagggcgt gcccctcgtg ggtctcgtgg tggaggggg ccctaacgtg gtgtccatcg 960 tcttggaata cctgcaagaa gagcctccca tccctgtggt gatttgtgat ggcagcggac gtgcctcgga catcctgtcc tttgcgcaca agtactgtga agaaggcgga ataataaatg 1020 1080 agtccctcag ggagcagctt ctagttacca ttcagaaaac atttaattat aataaggcac 1140 aatcacatca gctgtttgca attataatgg agtgcatgaa gaagaaagaa ctcgtcactg tgttcagaat gggttctgag ggccagcagg acatcgagat ggcaatttta actgccctgc 1200 1260 tgaaaggaac aaacgtatct gctccagatc agctgagctt ggcactggct tggaaccgcg tggacatagc acgaagccag atctttgtct ttgggcccca ctggacgccc ctgggaagcc 1320 tggcaccccc gacggacagc aaagccacgg agaaggagaa gaagccaccc atggccacca 1380 ccaagggagg aagaggaaaa gggaaaggca agaagaaagg gaaagtgaaa gaggaagtgg 1440 1500 aggaagaaac tgaccccgg aagatagagc tgctgaactg ggtgaatgct ttggagcaag 1560 cgatgctaga tgctttagtc ttagatcgtg tcgactttgt gaagctcctg attgaaaacg 1620 gagtgaacat gcaacacttt ctgaccattc cgaggctgga ggagctctat aacacaagac 1680 tgggtccacc aaacacactt catctgctgg tgagggatgt gaaaaagagc aaccttccgc ctgattacca catcagcctc atagacatcg ggctcgtgct ggagtacctc atgggaggag 1740 1800 cctaccgctg caactacact cggaaaaact ttcggaccct ttacaacaac ttgtttggac caaagaggcc taaagctctt aaacttctgg gaatggaaga tgatgagcct ccagctaaag 1860 1920 ggaagaaaaa aaaaaaaaag aaaaaggagg aagagatcga cattgatgtg gacgaccctg ccqtqaqtcq qttccagtat cccttccacg agctgatggt gtgggcagtg ctgatgaaac 1980 2040 gccagaaaat ggcagtgttc ctctggcagc gaggggaaga gagcatggcc aaggccctgg tggcctgcaa gctctacaag gccatggccc acgagtcctc cgagagtgat ctggtggatg 2100 2160 acatetecca ggaettggat aacaatteca aagaettegg ceagettget ttggagttat tagaccagtc ctataagcat gacgagcaga tcgctatgaa actcctgacc tacgagctga 2220 2280 aaaactggag caactcgacc tgcctcaaac tggccgtggc agccaaacac cgggacttca 2340 ttgctcacac ctgcagccag atgctgctga ccgatatgtg gatgggaaga ctgcggatgc 2400 ggaagaaccc cggcctgaag gttatcatgg ggattcttct accccccacc atcttgtttt 2460 tggaatttcg cacatatgat gatttctcgt atcaaacatc caaggaaaac gaggatggca aagaaaaaga agaggaaaat acggatgcaa atgcagatgc tggctcaaga aagggggatg 2520 2580 aggagaacga gcataaaaaa cagagaagta ttcccatcgg aacaaagatc tgtgaattct 2640 ataacgcgcc cattgtcaag ttctggtttt acacaatatc atacttgggc tacctgctgc 2700 tgtttaacta cgtcatcctg gtgcggatgg atggctggcc gtccctccag gagtggatcg tcatctccta catcgtgagc ctggcgttag agaagatacg agagatcctc atgtcagaac 2760 2820 caggcaaact cagccagaaa atcaaagttt ggcttcagga gtactggaac atcacagatc 2880 tegtggecat ttecacatte atgattggag caattetteg ectacagaac cageectaca

```
tgggctatgg ccgggtgatc tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg
                                                                      2940
acatctttgg tgtcaacaag tatctggggc catacgtgat gatgattgga aagatgatga
                                                                      3000
tegacatget gtactttgtg gteateatge tggtegtget catgagttte ggagtageee
                                                                      3060
                                                                      3120
gtcaagccat tetgcateca gaggagaage eetettggaa aetggeeega aacatettet
acatgcccta ctggatgatc tatggagagg tgtttgcaga ccagatagac ctctacgcca
                                                                      3180
tggaaattaa teeteettgt ggtgagaace tatatgatga ggagggeaag eggetteete
                                                                       3240
 cctgtatccc cggcgcctgg ctcactccag cactcatggc gtgctatcta ctggtcgcca
                                                                       3300
acatcctgct ggtgaacctg ctgattgctg tgttcaacaa tactttcttt gaagtaaaat
                                                                      3360
caatatccaa ccaggtgtgg aagttccagc gatatcagct gattatgaca tttcatgaca
                                                                      3420
 ggccagtcct gccccaccg atgatcattt taagccacat ctacatcatc attatgcgtc
                                                                       3480
 tcagcggccg ctgcaggaaa aagagagaag gggaccaaga ggaacgggat cgtggattga
                                                                       3540
agetetteet tagegaegag gagetaaaga ggetgeatga gttegaggag eagtgegtge
                                                                      3600
aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcatccggg
                                                                      3660
tcacttctga aagagttgaa aatatgtcaa tgaggttgga agaaatcaat gaaagagaaa
                                                                      3720
 cttttatgaa aacttccctg cagactgttg accttcgact tgctcagcta gaagaattat
                                                                       3780
ctaacagaat ggtgaatgct cttgaaaatc ttgcgggaat cgacaggtct gacctgatcc
                                                                      3840
 aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccgg caaagcagca
                                                                       3900
tcaatagcgc tgatggctac agcttgtatc gatatcattt taacggagaa gagttattat
                                                                      3960
ttgaggatac atctctctcc acgtcaccag ggacaggagt caggaaaaaa acctgttcct
                                                                      4020
                                                                      4080
tccgtataaa ggaagagaag gacgtgaaaa cgcacctagt cccagaatgt cagaacagtc
ttcacctttc actgggcaca agcacatcag caaccccaga tggcagtcac cttgcagtag
                                                                      4140
atgacttaaa gaacgctgaa gagtcaaaat taggtccaga tattgggatt tcaaaggaag
                                                                      4200
atgatgaaag acagacagac tctaaaaaag aagaaactat ttccccaagt ttaaataaaa
                                                                      4260
cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa
                                                                      4320
cgacaaatat agaaggcact atttcctatc ccctggaaga aaccaaaatt acacgctatt
                                                                      4380
tccccgatga aacgatcaat gcttgtaaaa caatgaagtc cagaagcttc gtctattccc
                                                                      4440
ggggaagaaa gctggtcggt ggggttaacc aggatgtaga gtacagttca atcacggacc
                                                                      4500
agcaattgac gacggaatgg caatgccaag ttcaaaagat cacgcgctct catagcacag
                                                                      4560
atatteetta cattgtgteg gaagetgeag tgeaagetga geaaaaagag cagtttgeag
                                                                      4620
atatgcaaga tgaacaccat gtcgctgaag caattcctcg aatccctcgc ttgtccctaa
                                                                      4680
ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg
                                                                      4740
gattcccatc tctcaggtca aaaagtttac atggacatcc taggaatgtg aaatccattc
                                                                      4800
agggaaagtt agacagatct ggacatgcca gtagtgtaag cagcttagta attgtgtctg
                                                                      4860
gaatgacagc agaagaaaaa aaggttaaga aagagaaagc ttccacagaa actgaatgct
                                                                      4920
                                                                      4980
agtctgtttt gtttctttaa ttttttttt taacagtcag aaacccacta atgggtgtca
tettggecca tectaaacae atmteeaatt teetaaaaae atttteeett aaaaaatttt
                                                                      5040
ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatatgt
                                                                      5100
tagtaggett agttttttca gttgcagtag tatcaaatga aagtgatgat actgtaacga
                                                                      5160
agataaattg gctaatcagt atacaagatt atacaatctc tttattactg agggccacca
                                                                      5220
aatagcctag gaagtgccct cgagcactga agtcaccatt aggtcactca agaagtaagc
                                                                      5280
aactagctgg gcacagtggc tcatgcctgt aatcctagca ctttgggagg ccaaggcaga
                                                                      5340
aagatagett gagteeagga gtttgagace ageetgggea acatagtgat accecatete
                                                                      5400
                                                                      5433
ttaaaaaaa aaaaaaaaa ctgccctcgt gcc
```

<210> 10

<211> 1533

<212> PRT

<213> Mus Musculus

<400> 10

Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His 10 Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser 20 25 30 Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln 35 40 Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp -10-

11e 65	Phe	Thr	Gly	Gly	Val 70	Ser	Thr	Gly	Val	Ile 75	Ser	His	Val	Gly	Asp 80
Ala	Leu	Lys	Asp	His 85	Ser	Ser	Lys	Ser	Arg 90	Gly	Arg	Val	Cys	Ala 95	Ile
Gly	Ile	Ala	Pro 100	Trp	Gly	Ile	Val	Glu 105	Asn	Lys	Glu	Asp	Leu 110	Val	Gly
Lys	Asp	Val 115	Thr	Arg	Val	Tyr	Gln 120	Thr	Met	Ser	Asn	Pro 125	Leu	Ser	Lys
Leu	Ser 130	Val	Leu	Asn	Asn	Ser 135	His	Thr	His	Phe	Ile 140	Leu	Ala	Asp	Asn
145			-	-	150	_				155				Leu	160
	_			165					170					Gln 175	
			180	-				185	_	-			190	Val	
		195					200					205		Val	
	210					215					220			His	
225	_				230					235				Gln	240
				245					250					Ser 255	
			260					265					270	Leu Met	
		275					280					285		Asp	
	290	Leu				295					300				
															tal n
305					310	-		_		315			_		320
305 Ile	Phe	Val	Phe	Gly 325	310 Pro	His	Trp	Thr	Pro 330	315 Leu	Gly	Ser	Leu	Ala 335	320 Pro
305 Ile Pro	Phe Thr	Val Asp	Phe Ser 340	Gly 325 Lys	310 Pro Ala	His Thr	Trp Glu	Thr Lys 345	Pro 330 Glu	315 Leu Lys	Gly Lys	Ser Pro	Leu Pro 350	Ala 335 Met	320 Pro Ala
305 Ile Pro Thr	Phe Thr Thr	Val Asp Lys 355	Phe Ser 340 Gly	Gly 325 Lys Gly	310 Pro Ala Arg	His Thr Gly	Trp Glu Lys 360	Thr Lys 345 Gly	Pro 330 Glu Lys	315 Leu Lys Gly	Gly Lys Lys	Ser Pro Lys 365	Leu Pro 350 Lys	Ala 335 Met Gly	320 Pro Ala Lys
305 Ile Pro Thr Val	Phe Thr Thr Lys 370	Val Asp Lys 355 Glu	Phe Ser 340 Gly Glu	Gly 325 Lys Gly Val	310 Pro Ala Arg Glu	His Thr Gly Glu 375	Trp Glu Lys 360 Glu	Thr Lys 345 Gly Thr	Pro 330 Glu Lys Asp	315 Leu Lys Gly Pro	Gly Lys Lys Arg 380	Ser Pro Lys 365 Lys	Leu Pro 350 Lys Ile	Ala 335 Met Gly Glu	320 Pro Ala Lys Leu
305 Ile Pro Thr Val Leu 385	Phe Thr Thr Lys 370 Asn	Val Asp Lys 355 Glu Trp	Phe Ser 340 Gly Glu Val	Gly 325 Lys Gly Val	310 Pro Ala Arg Glu Ala 390	His Thr Gly Glu 375 Leu	Trp Glu Lys 360 Glu Glu	Thr Lys 345 Gly Thr	Pro 330 Glu Lys Asp	315 Leu Lys Gly Pro Met 395	Gly Lys Lys Arg 380 Leu	Ser Pro Lys 365 Lys Asp	Leu Pro 350 Lys Ile Ala	Ala 335 Met Gly Glu Leu	320 Pro Ala Lys Leu Val 400
305 Ile Pro Thr Val Leu 385 Leu	Phe Thr Thr Lys 370 Asn	Val Asp Lys 355 Glu Trp Arg	Phe Ser 340 Gly Glu Val	Gly 325 Lys Gly Val Asn Asp 405	310 Pro Ala Arg Glu Ala 390 Phe	His Thr Gly Glu 375 Leu Val	Trp Glu Lys 360 Glu Glu Lys	Thr Lys 345 Gly Thr Gln Leu	Pro 330 Glu Lys Asp Ala Leu 410	Lys Gly Pro Met 395 Ile	Gly Lys Lys Arg 380 Leu Glu	Ser Pro Lys 365 Lys Asp	Leu Pro 350 Lys Ile Ala Gly	Ala 335 Met Gly Glu Leu Val 415	320 Pro Ala Lys Leu Val 400 Asn
305 Ile Pro Thr Val Leu 385 Leu	Phe Thr Thr Lys 370 Asn Asp	Val Asp Lys 355 Glu Trp Arg	Phe Ser 340 Gly Glu Val Val Phe 420	Gly 325 Lys Gly Val Asn Asp 405 Leu	310 Pro Ala Arg Glu Ala 390 Phe Thr	His Thr Gly Glu 375 Leu Val	Trp Glu Lys 360 Glu Glu Lys Pro	Thr Lys 345 Gly Thr Gln Leu Arg 425	Pro 330 Glu Lys Asp Ala Leu 410 Leu	Lys Gly Pro Met 395 Ile Glu	Gly Lys Lys Arg 380 Leu Glu	Ser Pro Lys 365 Lys Asp Asn Leu	Leu Pro 350 Lys Ile Ala Gly Tyr 430	Ala 335 Met Gly Glu Leu Val 415 Asn	320 Pro Ala Lys Leu Val 400 Asn
305 Ile Pro Thr Val Leu 385 Leu Met	Phe Thr Thr Lys 370 Asn Asp Gln Leu	Val Asp Lys 355 Glu Trp Arg His	Phe Ser 340 Gly Glu Val Val Phe 420 Pro	Gly 325 Lys Gly Val Asn Asp 405 Leu	310 Pro Ala Arg Glu Ala 390 Phe Thr	His Thr Gly Glu 375 Leu Val Ile	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440	Thr Lys 345 Gly Thr Gln Leu Arg 425 His	Pro 330 Glu Lys Asp Ala Leu 410 Leu	Lys Gly Pro Met 395 Ile Glu Leu	Gly Lys Lys Arg 380 Leu Glu Glu Val	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp	Ala 335 Met Gly Glu Leu Val 415 Asn	320 Pro Ala Lys Leu Val 400 Asn Thr
305 Ile Pro Thr Val Leu 385 Leu Met Arg	Phe Thr Thr Lys 370 Asn Asp Gln Leu Ser 450	Val Asp Lys 355 Glu Trp Arg His Gly 435 Asn	Phe Ser 340 Gly Glu Val Val Phe 420 Pro Leu	Gly 325 Lys Gly Val Asn Asp 405 Leu Pro	310 Pro Ala Arg Glu Ala 390 Phe Thr Asn	His Thr Gly Glu 375 Leu Val Ile Thr Asp 455	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440 Tyr	Thr Lys 345 Gly Thr Gln Leu Arg 425 His	Pro 330 Glu Lys Asp Ala Leu 410 Leu	Leu Lys Gly Pro Met 395 Ile Glu Leu Ser	Gly Lys Lys Arg 380 Leu Glu Glu Val Leu 460	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445 Ile	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp	Ala 335 Met Gly Glu Leu Val 415 Asn Val	320 Pro Ala Lys Leu Val 400 Asn Thr Lys
305 Ile Pro Thr Val Leu 385 Leu Met Arg Lys Leu 465	Phe Thr Thr Lys 370 Asn Asp Gln Leu Ser 450 Val	Val Asp Lys 355 Glu Trp Arg His Gly 435 Asn Leu	Phe Ser 340 Gly Glu Val Val Phe 420 Pro Leu Glu	Gly 325 Lys Gly Val Asn Asp 405 Leu Pro Pro	310 Pro Ala Arg Glu Ala 390 Phe Thr Asn Pro Leu 470	His Thr Gly Glu 375 Leu Val Ile Thr Asp 455 Met	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440 Tyr	Thr Lys 345 Gly Thr Gln Leu Arg 425 His Gly	Pro 330 Glu Lys Asp Ala Leu 410 Leu Leu	Jis Leu Lys Gly Pro Met Jis Glu Leu Ser Tyr	Gly Lys Arg 380 Leu Glu Glu Val Leu 460 Arg	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445 Ile Cys	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp Asp	Ala 335 Met Gly Glu Leu Val 415 Asn Val Ile	320 Pro Ala Lys Leu Val 400 Asn Thr Lys Gly Thr 480
305 Ile Pro Thr Val Leu 385 Leu Met Arg Lys Leu 465 Arg	Phe Thr Thr Lys 370 Asn Asp Gln Leu Ser 450 Val Lys	Val Asp Lys 355 Glu Trp Arg His Gly 435 Asn Leu Asn	Phe Ser 340 Gly Glu Val Val Phe 420 Pro Leu Glu Phe	Gly 325 Lys Gly Val Asn Asp 405 Leu Pro Pro Tyr Arg 485	310 Pro Ala Arg Glu Ala 390 Phe Thr Asn Pro Leu 470 Thr	His Thr Gly Glu 375 Leu Val Ile Thr Asp 455 Met Leu	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440 Tyr Gly	Thr Lys 345 Gly Thr Gln Leu Arg 425 His Gly Asn	Pro 330 Glu Lys Asp Ala Leu 410 Leu Ile Ala Asn 490	Lys Gly Pro Met 395 Ile Glu Leu Ser Tyr 475 Leu	Gly Lys Arg 380 Leu Glu Glu Val Leu 460 Arg	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445 Ile Cys	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp Asp Asp	Ala 335 Met Gly Glu Leu Val 415 Asn Val Ile Tyr Lys 495	320 Pro Ala Lys Leu Val 400 Asn Thr Lys Gly Thr 480 Arg
305 Ile Pro Thr Val Leu 385 Leu Met Arg Lys Leu 465 Arg	Phe Thr Thr Lys 370 Asn Asp Gln Leu Ser 450 Val Lys Lys	Val Asp Lys 355 Glu Trp Arg His Gly 435 Asn Leu Asn	Phe Ser 340 Gly Glu Val Val Phe 420 Pro Leu Glu Phe Leu 500	Gly 325 Lys Gly Val Asn Asp 405 Leu Pro Pro Tyr Arg 485 Lys	310 Pro Ala Arg Glu Ala 390 Phe Thr Asn Pro Leu 470 Thr	His Thr Gly Glu 375 Leu Val Ile Thr Asp 455 Met Leu	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440 Tyr Gly Tyr	Thr Lys 345 Gly Thr Gln Leu Arg 425 His Gly Asn Met 505	Pro 330 Glu Lys Asp Ala Leu 410 Leu Ile Ala Asn 490 Glu	Lys Gly Pro Met 395 Ile Glu Leu Ser Tyr 475 Leu Asp	Gly Lys Arg 380 Leu Glu Glu Val Leu 460 Arg Phe	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445 Ile Cys Gly	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp Asp Asn Pro Pro 510	Ala 335 Met Gly Glu Leu Val 415 Asn Val Ile Tyr Lys 495 Pro	320 Pro Ala Lys Leu Val 400 Asn Thr Lys Gly Thr 480 Arg
305 Ile Pro Thr Val Leu 385 Leu Met Arg Lys Leu 465 Arg Pro	Phe Thr Thr Lys 370 Asn Asp Gln Leu Ser 450 Val Lys Lys Gly	Val Asp Lys 355 Glu Trp Arg His Gly 435 Asn Leu Asn Ala Lys 515	Phe Ser 340 Gly Glu Val Val Phe 420 Pro Leu Glu Phe Leu 500 Lys	Gly 325 Lys Gly Val Asn Asp 405 Leu Pro Pro Tyr Arg 485 Lys	310 Pro Ala Arg Glu Ala 390 Phe Thr Asn Pro Leu 470 Thr Leu Lys	His Thr Gly Glu 375 Leu Val Ile Thr Asp 455 Met Leu Leu	Trp Glu Lys 360 Glu Glu Lys Pro Leu 440 Tyr Gly Tyr Gly Lys 520	Thr Lys 345 Gly Thr Gln Leu Arg 425 His Gly Asn Met 505 Lys	Pro 330 Glu Lys Asp Ala Leu 410 Leu Ile Ala Asn 490 Glu Lys	Jis Leu Lys Gly Pro Met Jis Glu Leu Ser Tyr 475 Leu Asp	Gly Lys Arg 380 Leu Glu Glu Val Leu 460 Arg Phe Asp Glu	Ser Pro Lys 365 Lys Asp Asn Leu Arg 445 Ile Cys Gly Glu Glu 525	Leu Pro 350 Lys Ile Ala Gly Tyr 430 Asp Asp Pro Pro 510 Ile	Ala 335 Met Gly Glu Leu Val 415 Asn Val Ile Tyr Lys 495	320 Pro Ala Lys Leu Val 400 Asn Thr Lys Gly Thr 480 Arg Ala

-11-Leu Met Val Trp Ala Val Leu Met Lys Arg Gln Lys Met Ala Val Phe · 550 Leu Trp Gln Arg Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys Leu Tyr Lys Ala Met Ala His Glu Ser Ser Glu Ser Asp Leu Val Asp Asp Ile Ser Gln Asp Leu Asp Asn Asn Ser Lys Asp Phe Gly Gln Leu Ala Leu Glu Leu Leu Asp Gln Ser Tyr Lys His Asp Glu Gln Ile Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys Leu Lys Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His Thr Cys Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg Met Arg Lys Asn Pro Gly Leu Lys Val Ile Met Gly Ile Leu Leu Pro Pro Thr Ile Leu Phe Leu Glu Phe Arg Thr Tyr Asp Asp Phe Ser Tyr 690 700 Gln Thr Ser Lys Glu Asn Glu Asp Gly Lys Glu Lys Glu Glu Glu Asn Thr Asp Ala Asn Ala Asp Ala Gly Ser Arg Lys Gly Asp Glu Glu Asn Glu His Lys Lys Gln Arg Ser Ile Pro Ile Gly Thr Lys Ile Cys Glu
740 745 750 Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Ile Ser Tyr Leu Gly Tyr Leu Leu Phe Asn Tyr Val Ile Leu Val Arg Met Asp Gly Trp Pro Ser Leu Gln Glu Trp Ile Val Ile Ser Tyr Ile Val Ser Leu Ala Leu Glu Lys Ile Arg Glu Ile Leu Met Ser Glu Pro Gly Lys Leu Ser Gln Lys Ile Lys Val Trp Leu Gln Glu Tyr Trp Asn Ile Thr Asp Leu Val Ala Ile Ser Thr Phe Met Ile Gly Ala Ile Leu Arg Leu Gln Asn Gln Pro Tyr Met Gly Tyr Gly Arg Val Ile Tyr Cys Val Asp Ile Ile Phe Trp Tyr Ile Arg Val Leu Asp Ile Phe Gly Val Asn Lys Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met Leu Tyr Phe Val Val Ile Met Leu Val Val Leu Met Ser Phe Gly Val Ala Arg Gln Ala Ile Leu His Pro Glu Glu Lys Pro Ser Trp Lys Leu Ala Arg Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu Val Phe Ala Asp Gln Ile Asp Leu Tyr Ala Met Glu Ile Asn Pro Pro Cys Gly Glu Asn Leu Tyr Asp Glu Glu Gly Lys Arg Leu Pro Pro Cys Ile Pro Gly Ala Trp Leu Thr Pro Ala Leu Met Ala Cys Tyr Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Val Phe Asn Asn Thr Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val Trp Lys Phe Gln Arg

Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro · 1030 1035 Met Ile Ile Leu Ser His Ile Tyr Ile Ile Ile Met Arg Leu Ser Gly 1045 1050 1055 Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly 1060 1065 1070 Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe 1075 1080 1085 Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln 1090 1095 1100 Gln Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu 1110 1115 1105 112 Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met 1125 1130 1135 Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu 1140 1145 1150 Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp 1155 1160 1165 Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu 1170 1175 1180 Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr 1190 1195 Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp 1205 1210 1215 Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys 1220 1225 1230 Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro 1235 1240 1245 Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala 1250 1255 1260 Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu 1270 1275 128 Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu 1285 1290 1295 Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn 1300 1305 1310 Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr 1315 1320 1325 Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro 1330 1335 1340 Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn 1345 1350 1355 136 Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg 1365 1370 1375 Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr 1380 1385 1390 Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr 1395 1400 1405 Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val 1410 1415 1420 Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His 1430 1435 1425 Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr 1445 1450 1455 Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr 1460 1465 1470 Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg 1475 1480 1485 Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser 1495 1500

-13-

 Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys

 1505
 : 1510
 1515
 152

 Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
 1525
 1530

<210> 11 <211> 6220 <212> DNA <213> Homo Sapiens

<400> 11

tgtgcagaat tgtacagttg cgaaaccatg tcgctggcag ctggtgctgg cggtggagac 60 ttecctgtgc ggtgctcagt gcatctgcac ccgtggggga gggagctctt tctctggccc 120 180 tgcagtcacc tgaggttgtt accattatga acggccgctg ggacccccgc atgtgcatgt actocccag agtgtccggg ggccccagcc aagggacaca totcacgcag ctgggaacat 240 300 gtgcaggctg atgaagagaa ccggatgagg gcttcacatg aggaagcatg tggccaggtc 360 ctctcagaac atcagcctca tcttcctgtc tctgatctat ttcaccaacc accccatgtg 420 tctctagaac cccagtgtag cgagctggag agaggactgt cctgagggca gcaggcctgg ttqcaqctqq cqtqqqqtc tcaqaatgga gccctcagcc ctgaggaaag ctggctcgga 480 540 gcaggaggag ggctttgagg ggctgcccag aagggtcact gacctgggga tggtctccaa totocogogo agcaacagoa gootottoaa gagotggagg otacagtgoo cottoggoaa 600 caatgacaag caagaaagcc tcagttcgtg gattcctgaa aacatcaaga agaaagaatg 660 cgtgtatttt gtggaaagtt ccaaactgtc tgatgctggg aaggtggtgt gtcagtgtgg 720 ctacacgcat gagcagcact tggaggaggc taccaagccc cacaccttcc agggcacaca 780 840 gtgggaccca aagaaacatg tccaggagat gccaaccgat gcctttggcg acatcgtctt 900 cacgggcctg agccagaagg tgaaaaagta cgtccgagtc tcccaggaca cgccctccag 960 cgtgatctac cacctcatga cccagcactg ggggctggac gtccccaatc tcttgatctc 1020 ggtgaccggg ggggccaaga acttcaacat gaagccgcgg ctgaagagca ttttccgcag 1080 aggeotggte aaggtggete agaccacagg ggcotggate atcacagggg ggtoccacac 1140 cggcgtcatg aagcaggtag gcgaggcggt gcgggacttc agcctgagca gcagctacaa qqaaqqcqaq ctcatcacca tcqqaqtcqc cacctqqqqc actqtccacc gccgcqaggg 1200 1260 cctgatccat cccacggca gcttccccgc cgagtacata ctggatgagg atggccaagg gaacctgacc tgcctagaca gcaaccactc tcacttcatc ctcgtggacg acgggaccca 1320 1380 cggccagtac ggggtggaga ttcctctgag gaccaggctg gagaagttca tatcggagca gaccaaggaa agaggaggtg tggccatcaa gatccccatc gtgtgcgtgg tgctggaggg 1440 1500 cggcccgggc acgttgcaca ccatcgacaa cgccaccacc aacggcaccc cctgtgtggt tgtggagggc tcgggccgcg tggccgacgt cattgcccag gtggccaacc tgcctgtctc 1560 ggacatcact atctccctga tccagcagaa actgagcgtg ttcttccagg agatgtttga 1620 gaccttcacg gaaagcagga ttgtcgagtg gaccaaaaag atccaagata ttgtccggag 1680 1740 gcggcagctg ctgactgtct tccgggaagg caaggatggt cagcaggacg tggatgtggc 1800 catcttgcag gccttgctga aagcctcacg gagccaagac cactttggcc acgagaactg ggaccaccag ctgaaactgg cagtggcatg gaatcgcgtg gacattgccc gcagtgagat 1860 cttcatggat gagtggcagt ggaagccttc agatctgcac cccacgatga cagctgcact 1920 catctccaac aagcctgagt ttgtgaagct cttcctggaa aacggggtgc agctgaagga 1980 gtttgtcacc tgggacacct tgctctacct gtacgagaac ctggacccct cctgcctgtt 2040 ccacagcaag ctgcaaaagg tgctggtgga ggatcccgag cgcccggctt gcgcgcccgc 2100 ggcgccccgc ctgcagatgc accacgtggc ccaggtgctg cgggagctgc tgggggactt 2160 2220 cacgcagccg ctttatcccc ggccccggca caacgaccgg ctgcggctcc tgctgcccgt tecceacgte aageteaacg tgcagggagt gageeteegg tecetetaca agegtteete 2280 aggccatgtg accttcacca tggaccccat ccgtgacctt ctcatttggg ccattgtcca 2340 2400 gaaccgtcgg gagctggcag gaatcatctg ggctcagagc caggactgca tcgcagcggc 2460 cttggcctgc agcaagatcc tgaaggaact gtccaaggag gaggaggaca cggacagctc 2520 ggaggagatg ctggcgctgg cggaggagta tgagcacaga gccatcgggg tcttcaccga gtgctaccgg aaggacgaag agagagccca gaaactgctc acccgcgtgt ccgaggcctg 2580 2640 ggggaagacc acctgcctgc agctcgccct ggaggccaag gacatgaagt ttgtgtctca 2700 cgggggcatc caggccttcc tgaccaaggt gtggtggggc cagctctccg tggacaatgg 2760 getgtggcgt gtgaccetgt geatgetgge ettecegetg etecteaceg geetcatete 2820 cttcagggag aagaggctgc aggatgtggg cacccccgcg gcccgcgccc gtgccttctt caccqcaccc gtggtggtct tccacctgaa catcctctcc tacttcgcct tcctctgcct 2880 2940 gttcgcctac gtgctcatgg tggacttcca gcctgtgccc tcctggtgcg agtgtgccat

-14-

			-14-			
ctacctctgg	ctcttctcct	tggtgtgcga	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gtacttcagt	gacttctgga	ataagctgga	3060
	atcttgctct					3120
	cgcgtcatcc					3180
	atcagtaaga					3240
	ttcttcctct					3300
	ctcatccaca					3360
	ctcaccatct					3420
	tgcagcccca					3480
	cagaggccgg					3540
	accaacatcc					3600
	caggagcaca					3660
	ggccgccccg					3720
	agggtggtcc					3780
	aacgaggagg					3840
cctccagaag	cgacagttcc	aggeeeegee	acceegggag	cadaadatcd	aggagaacta	3900
	gacgccatgg					3960
						4020
	aggttggcct aggacgctgc					4080
						4140
	aaggccgcgg					4200
	gacagctacc					4260
	cccgtgccca					4320
	tacacggcag					4380
	ctgtccacga					4440
	ccgtacacag					4500
	cgcgggagcc					4560
	aggcggaacg					4620
	gtggtgaagc					4680
	gagatgctac					4740
gtettttgaa	aacttgctga	agtgeggeat	ggaggtgtat	aaaggctaca	togatgaccc	4800
	gacaatgcct					4860
	gagctgaaca					4920
	caggtggtgg					4980
	gccgctgagt					5040
	gacgttcccc					5100
tcaggctgtt	cctgggccct	gcacatgatg	gggtttggtg	gacccagtgc	ccccaegge	
	ctgctgcaga					5160 5220
	acagggcaca					
	gctgggcctg					5280
	cgagcagctc					5340
	tcccctgtgg					5400 5460
	agggcctcca					
	ctgtctgaac					5520
	ccaggcctgt					5580
	gaggctgcag					5640
	ggacggcatg					5700
	aggagcgtga					5760
	ctgccatctc					5820
	ccagggccag					5880
	ctgcagcctg					5940
cctctgaccc	atcagctccc	tgtcttcatt	ttcctaaacc	tgggctccag	catcgtcccc	6000
	ggccaggatg					6060
	tgtgccctgg					6120
	actggccgag			gacaggcctc	tgtcttccac	6180
ctgacccaaa	gctctctagc	cacccccttg	tccccagtat			6220
					•	

<212> PRT <213> Homo Sapiens

<400> 12 Met Glu Pro Ser Ala Leu Arg Lys Ala Gly Ser Glu Gln Glu Gly Phe Glu Gly Leu Pro Arg Arg Val Thr Asp Leu Gly Met Val Ser Asn Leu Arg Arg Ser Asn Ser Ser Leu Phe Lys Ser Trp Arg Leu Gln Cys Pro Phe Gly Asn Asn Asp Lys Gln Glu Ser Leu Ser Ser Trp Ile Pro Glu Asn Ile Lys Lys Glu Cys Val Tyr Phe Val Glu Ser Ser Lys Leu Ser Asp Ala Gly Lys Val Val Cys Gln Cys Gly Tyr Thr His Glu Gln His Leu Glu Glu Ala Thr Lys Pro His Thr Phe Gln Gly Thr Gln Trp Asp Pro Lys Lys His Val Gln Glu Met Pro Thr Asp Ala Phe Gly Asp Ile Val Phe Thr Gly Leu Ser Gln Lys Val Lys Lys Tyr Val Arg Val Ser Gln Asp Thr Pro Ser Ser Val Ile Tyr His Leu Met Thr Gln His Trp Gly Leu Asp Val Pro Asn Leu Leu Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Asn Met Lys Pro Arg Leu Lys Ser Ile Phe Arg Arg Gly Leu Val Lys Val Ala Gln Thr Thr Gly Ala Trp Ile Ile Thr Gly Gly Ser His Thr Gly Val Met Lys Gln Val Gly Glu Ala Val Arg Asp Phe Ser Leu Ser Ser Tyr Lys Glu Gly Glu Leu Ile Thr Ile Gly Val Ala Thr Trp Gly Thr Val His Arg Arg Glu Gly Leu Ile His Pro 245 250 255 Thr Gly Ser Phe Pro Ala Glu Tyr Ile Leu Asp Glu Asp Gly Gln Gly Asn Leu Thr Cys Leu Asp Ser Asn His Ser His Phe Ile Leu Val Asp Asp Gly Thr His Gly Gln Tyr Gly Val Glu Ile Pro Leu Arg Thr Arg Leu Glu Lys Phe Ile Ser Glu Gln Thr Lys Glu Arg Gly Gly Val Ala 305 310 315 320Ile Lys Ile Pro Ile Val Cys Val Val Leu Glu Gly Gly Pro Gly Thr Leu His Thr Ile Asp Asn Ala Thr Thr Asn Gly Thr Pro Cys Val Val Val Glu Gly Ser Gly Arg Val Ala Asp Val Ile Ala Gln Val Ala Asn Leu Pro Val Ser Asp Ile Thr Ile Ser Leu Ile Gln Gln Lys Leu Ser Val Phe Phe Gln Glu Met Phe Glu Thr Phe Thr Glu Ser Arg Ile Val Glu Trp Thr Lys Lys Ile Gln Asp Ile Val Arg Arg Arg Gln Leu Leu Thr Val Phe Arg Glu Gly Lys Asp Gly Gln Gln Asp Val Asp Val Ala Ile Leu Gln Ala Leu Leu Lys Ala Ser Arg Ser Gln Asp His Phe Gly

-16-

His	Glu 450	Asn	Trp	Asp	His	Gln 455	Leu	Lys	Leu	Ala	Val 460	Ala	Trp	Asn	Arg
Val 465		Ile	Ala	Arg	Ser 470	Glu	Ile	Phe	Met	Asp 475	Glu	Trp	Gln	Trp	Lys 480
	Ser	Asp	Leu			Thr	Met	Thr			Leu	Ile	Ser	Asn	
Pro	Glu	Phe	Val	485 Lys	Leu	Phe	Leu	Glu	490 Asn	Gly	Val	Gln	Leu	495 Lys	Glu
Phe	Val	Thr	500 Trp	Asp	Thr	Leu	Leu	505 Tvr	Leu	Tvr	Glu	Asn	510 Leu	Asp	Pro
		515	_	_			520			_		525		_	
	530					535			_		540			Asp	
Glu 545	Arg	Pro	Ala	Cys	Ala 550	Pro	Ala	Ala	Pro	Arg 555	Leu	Gln	Met	His	His 560
Val	Ala	Gln	Val	Leu 565	Arg	Glu	Leu	Leu	Gly 570	Asp	Phe	Thr	Gln	Pro 575	Leu
Tyr	Pro	Arg	Pro 580		His	Asn	Asp	Arģ 585	Leu	Arg	Leu	Leu	Leu 590	Pro	Val
Pro	His	Val 595		Leu	Asn	Val	Gln 600		Val	Ser	Leu	Arg 605		Leu	Tyr
Lys	Arg 610		Ser	Gly	His	Val 615		Phe	Thr	Met	Asp 620		Ile	Arg	Asp
Leu		Ile	Trp	Ala			Gln	Asn	Arg			Leu	Ala	Gly	
625	_			_	630	_	_	~ 1		635		_		_	640
ile	Trp	Ala	GIn	Ser 645	Gin	Asp	Cys	IIe	650	Ата	Ala	Leu	Ala	Cys 655	Ser
Lys	Ile	Leu	Lys 660	Glu	Leu	Ser	Lys	Glu 665	Glu	Glu	Asp	Thr	Asp 670	Ser	Ser
Glu	Glu	Met 675	Leu	Ala	Leu	Ala	Glu 680	Glu	Tyr	Glu	His	Arg 685	Ala	Ile	Gly
Val	Phe 690	Thr	Glu	Cys	Tyr	Arg 695	Lys	Asp	Glu	Glu	Arg 700	Ala	Gln	Lys	Leu
Leu 705	Thr	Arg	Val	Ser	Glu 710	Ala	Trp	Gly	Lys	Thr 715	Thr	Cys	Leu	Gln	Leu 720
Ala	Leu	Glu	Ala	Lys 725	Asp	Met	Lys	Phe	Val 730	Ser	His	Gly	Gly	Ile 735	Gln
Ala	Phe	Leu	Thr 740		Val	Trp	Trp	Gly 745		Leu	Ser	Val	Asp 750	Asn	Gly
Leu	Trp	Arg 755		Thr	Leu	Cys	Met 760		Ala	Phe	Pro	Leu 765		Leu	Thr
Gly			Ser	Phe	Arg			Arg	Leu	Gln	_		Gly	Thr	Pro
Ala	770 Ala	Arg	Ala	Arg	Ala	775 Phe	Phe	Thr	Ala	Pro	780 Val	Val	Val	Phe	His
785		T1 -	.	a	790	D1	n 1 -	n		795	.	Db -	71-	m	800
				805	-				810	_				Tyr 815	
Leu	Met	Val	Asp 820	Phe	Gln	Pro	Val	Pro 825	Ser	Trp	Cys	Glu	Cys 830	Ala	Ile
Tyr	Leu	Trp 835	Leu	Phe	Ser	Leu	Val 840	Cys	Glu	Glu	Met	Arg 845	Gln	Leu	Phe
Tyr	Asp 850	Pro	Asp	Glu	Cys	Gly 855	Leu	Met	Lys	Lys	Ala 860	Ala	Leu	Tyr	Phe
Ser 865		Phe	Trp	Asn	Lys 870	Leu	Asp	Val	Gly	Ala 875		Leu	Leu	Phe	Val 880
	Gly	Leu	Thr	-		Leu	Ile	Pro			Leu	Tyr	Pro	Gly	
Val	Ile	Leu		885 Leu	Asp	Phe	Ile		890 Phe	Cys	Leu	Arg		895 Met	His
Ile	Phe	Thr	900 Ile	Ser	Lys	Thr	Leu	905 Gly	Pro	Lys	Ile	Ile	910 Ile	Val	Lys
		915			_		920	•		-		925			-

-17-Arg Met Met Lys Asp Val Phe Phe Phe Leu Phe Leu Leu Ala Val Trp 930 935 940 Val Val Ser Phe Gly Val Ala Lys Gln Ala Ile Leu Ile His Asn Glu 955 950 945 Arg Arg Val Asp Trp Leu Phe Arg Gly Ala Val Tyr His Ser Tyr Leu 965 970 975 Thr Ile Phe Gly Gln Ile Pro Gly Tyr Ile Asp Gly Val Asn Phe Asn 980 985 990 Pro Glu His Cys Ser Pro Asn Gly Thr Asp Pro Tyr Lys Pro Lys Cys 1000 995 1005 Pro Glu Ser Asp Ala Thr Gln Gln Arg Pro Ala Phe Pro Glu Trp Leu 1010 1015 1020 Thr Val Leu Leu Cys Leu Tyr Leu Leu Phe Thr Asn Ile Leu Leu 1025 1030 1035 Leu Asn Leu Leu Ile Ala Met Phe Asn Tyr Thr Phe Gln Gln Val Gln 1045 1050 1055 Glu His Thr Asp Gln Ile Trp Lys Phe Gln Arg His Asp Leu Ile Glu 1060 1065 1070 Glu Tyr His Gly Arg Pro Ala Ala Pro Pro Pro Phe Ile Leu Leu Ser 1075 1080 1085 His Leu Gln Leu Phe Ile Lys Arg Val Val Leu Lys Thr Pro Ala Lys 1090 1095 1100Arg His Lys Gln Leu Lys Asn Lys Leu Glu Lys Asn Glu Glu Ala Ala 1105 1110 1115 112 Leu Leu Ser Trp Glu Ile Tyr Leu Lys Glu Asn Tyr Leu Gln Asn Arg 1125 1130 1135 Gln Phe Gln Gln Lys Gln Arg Pro Glu Gln Lys Ile Glu Asp Ile Ser 1140 1145 11501140 Asn Lys Val Asp Ala Met Val Asp Leu Leu Asp Leu Asp Pro Leu Lys 1155 1160 1165 Arg Ser Gly Ser Met Glu Gln Arg Leu Ala Ser Leu Glu Glu Gln Val 1180 1170 1175 Ala Gln Thr Ala Arg Ala Leu His Trp Ile Val Arg Thr Leu Arg Ala 1190 1195 Ser Gly Phe Ser Ser Glu Ala Asp Val Pro Thr Leu Ala Ser Gln Lys 1205 1210 1215 Ala Ala Glu Glu Pro Asp Ala Glu Pro Gly Gly Arg Lys Lys Thr Glu 1220 1225 1230 Glu Pro Gly Asp Ser Tyr His Val Asn Ala Arg His Leu Leu Tyr Pro 1245 1235 1240 Asn Cys Pro Val Thr Arg Phe Pro Val Pro Asn Glu Lys Val Pro Trp 1250 1255 1260 Glu Thr Glu Phe Leu Ile Tyr Asp Pro Pro Phe Tyr Thr Ala Glu Arg 1270 1275 1265 128 Lys Asp Ala Ala Ala Met Asp Pro Met Gly Asp Thr Leu Glu Pro Leu 1285 1290 1295 Ser Thr Ile Gln Tyr Asn Val Val Asp Gly Leu Arg Asp Arg Arg Ser 1300 1305 1310 Phe His Gly Pro Tyr Thr Val Gln Ala Gly Leu Pro Leu Asn Pro Met 1320 1315 1325 Gly Arg Thr Gly Leu Arg Gly Arg Gly Ser Leu Ser Cys Phe Gly Pro 1330 1335 1340 Asn His Thr Leu Tyr Pro Met Val Thr Arg Trp Arg Arg Asn Glu Asp 1350 1345 1355 136 Gly Ala Ile Cys Arg Lys Ser Ile Lys Lys Met Leu Glu Val Leu Val 1365 1370 1375 Val Lys Leu Pro Leu Ser Glu His Trp Ala Leu Pro Gly Gly Ser Arg 1380 1385 1390 Glu Pro Gly Glu Met Leu Pro Arg Lys Leu Lys Arg Ile Leu Arg Gln 1400 1405

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val
1410

Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile
1425

Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu
1445

Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser
1460

Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His
1475

Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr
1490

<210> 13 <211> 1816 <212> PRT

<213> C. Elegans

<400> 13

Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn 10 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser 25 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr 40 45 35 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu 55 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser 70 75 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu 85 90 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg 100 105 Glu Thr Glu Thr Gln Thr Lys Arg Arg Arg Lys Lys Gln Arg Ser Thr 120 125 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys 135 140 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys 155 Arg Gly Asn Ser Thr Thr Ser His Ser Gly His Ala Thr Arg Ala Gly 170 165 175 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln 185 Cys Ser Tyr Phe Val Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys 195 200 205 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln 215 220 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu 230 235 Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asn Glu Val Asn Leu Thr 245 250 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp 260 265 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr 275 280 285 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met 295 300 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn 305 310 315 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe

-19-Arg Lys Gly Ile Met Lys Ile Ala Lys Ser Thr Asp Ala Trp Ile Ile Thr Ser Gly Leu Asp Glu Gly Val Val Lys His Leu Asp Ser Ala Leu His Ala Leu Glu Phe Trp Ser Phe Gly Leu Phe Trp Val Ile Gln Leu Asp Val Leu Leu Ala His Ser Met Phe Ile Pro Arg Gly Ser Leu Phe Asp His Gly Asn His Thr Ser Lys Asn His Val Val Ala Ile Gly Ile Ala Ser Trp Gly Met Leu Lys Gln Arg Ser Arg Phe Val Gly Lys Asp Ser Thr Val Thr Tyr Ala Thr Asn Val Phe Asn Asn Thr Arg Leu Lys Glu Leu Asn Asp Asn His Ser Tyr Phe Leu Phe Ser Asp Asn Gly Thr Val Asn Arg Tyr Gly Ala Glu Ile Ile Met Arg Lys Arg Leu Glu Ala Tyr Leu Ala Gln Gly Asp Lys Lys Arg Ser Ala Ile Pro Leu Val Cys Val Val Leu Glu Gly Gly Ala Phe Thr Ile Lys Met Val His Asp Tyr Val Thr Thr Ile Pro Arg Ile Pro Val Ile Val Cys Asp Gly Ser Gly Arg Ala Ala Asp Ile Leu Ala Phe Ala His Gln Ala Val Ser Gln Asn Gly Phe Leu Ser Asp Asn Ile Arg Asn Gln Leu Val Asn Ile Val Arg Arg Ile Phe Gly Tyr Asp Pro Lys Thr Ala Gln Lys Leu Ile Lys Gln Ile Val Glu Cys Ser Thr Asn Lys Ser Leu Met Thr Ile Phe Arg Leu Gly Glu Ser Ser Arg Glu Asp Leu Asp His Val Ile Met Ser Cys Leu Leu Lys Gly Gln Asn Leu Ser Pro Pro Glu Gln Leu Gln Leu Ala Leu Ala Trp Asn Arg Ala Asp Ile Ala Arg Thr Glu Ile Phe Ala Asn Gly Thr Glu Trp Thr Thr Gln Asp Leu His Asn Ala Met Ile Glu Ala Leu Ser Asn Asp Arg Ile Asp Phe Val His Leu Leu Leu Glu Asn Gly Val Ser Met Gln Lys Phe Leu Thr Tyr Gly Arg Leu Glu His Leu Tyr Asn Thr Asp Lys Gly Pro Gln Asn Thr Leu Arg Thr Asn Leu Leu Val Asp · 700 Ser Lys His His Ile Lys Leu Val Glu Val Gly Arg Leu Val Glu Asn Leu Met Gly Asn Leu Tyr Lys Ser Asn Tyr Thr Lys Glu Glu Phe Lys Asn Gln Tyr Phe Leu Phe Asn Asn Arg Lys Gln Phe Gly Lys Arg Val His Ser Asn Ser Asn Gly Gly Arg Asn Asp Val Ile Gly Pro Ser Gly Asp Ala Gly Arg Glu Arg Met Ser Ser Met Gln Ile Ser Leu Ile Asn Asn Ala Arg Asn Ser Ile Ile Ser Leu Phe Asn Gly Gly Gly Arg Lys Arg Glu Ser Asp Asp Glu Asp Asp Phe Ser Asn Leu Glu Glu Ala -20-

				805					810					815	
		_	820				Tyr	825					830		
		835		•			Lys 840					845			
_	850					855	Ala				860				
865			-		870		Leu			875			_	•	880
	_			885			Asp		890					895	
			900				His	905					910		
		915					Trp 920	=	-			925			
	930					935	Lys				940				
945					950		Gln			955					960
Gln	Asn	Ser	Lys	Val 965	Leu	Thr	Cys	Leu	Ala 970	Ala	Pro	Pro	Leu	Ile 975	Phe
		_	980	_		_	Glu	985		•			990		
Ala	Ala	Glu 995	His	Asp	Glu	Glu	Met 1000		Asp	Ser	Glu	Met 1005		Ser	Ala
	1010)	_			1015					1020)			
1025	5				1030)	Arg			103	5				104
Pro	Leu	Ser	Ile	His 1045		Leu	Val	Arg	Asp 1050		Leu	Asn	Phe	Ser 1055	
			1060)			Ile	1069	5				1070)	
		1075	5				Ala 1080)				1085	5		
Lys	Lys 1090		Val	Ile	Lys	Pro 1095	Pro	Ala	Cys	Leu	Lys 1100		Glu	Thr	Ser
Asp 1105		Asp	Glu	Gln	Glu 1110		Lys	Lys	Ala	Thr 1119		Met	Cys	Lys	Ser 112
Thr	Phe	Phe	Asp	Phe 1125		Phe	Asp	Phe	Pro 1130	_	Ile	Asn	Arg	Thr 1135	_
Lys	Arg	Gly	Ser 1140		Ala	Val	Ala	Met 1145		His	Asp	Asp	Met 1150		Ile
Asp	Pro	Ser 1155		Glu	Leu	Asp	Thr 1160		Thr	Arg	Gln	Lys 1165		Ser	Arg
Glu	Phe 1170		Ser	Ser	Arg	Asn 1175	Val	Thr	Val	Gln	Val 1180	-	Thr	Gln	Arg
		Ser	Trp	Lys			Ile	Met	Glu			Lys	Ala	Pro	
1185		_	_	_	1190					119!			_	-1.	120
				1205	5		Phe		1210)				1215	5
		_	1220)			Lys	1225	5				1230)	
		1235	5				Ile 1240)				1245	5		
_	1250)				1255			_		1260)			
1265	5				1270)	Phe			127	5				128
Phe	Gly	Leu	Leu	Thr	Tyr	Leu	Ile	Ala	Tyr	Phe	Ile	Arg	Leu	Ser	Pro

-21-

1285 1290 1295 Thr Thr Lys Thr Leu Gly Arg Ile Leu Ile Ile Cys Asn Ser Val Ile 1300 1305 1310 Trp Ser Leu Lys Leu Val Asp Tyr Leu Ser Val Gln Gln Gly Leu Gly 1315 1320 1325 Pro Tyr Ile Asn Ile Val Ala Glu Met Ile Pro Thr Met Ile Pro Leu 1330 1335 1340 Cys Val Leu Val Phe Ile Thr Leu Tyr Ala Phe Gly Leu Leu Arg Gln 1350 1355 1345 Ser Ile Thr Tyr Pro Tyr Glu Asp Trp His Trp Ile Leu Val Arg Asn 1365 1370 1375 Ile Phe Leu Gln Pro Tyr Phe Met Leu Tyr Gly Glu Val Tyr Ala Ala 1380 1385 1390 Glu Ile Asp Thr Cys Gly Asp Glu Ile Trp Gln Thr His Glu Asp Glu 1395 1400 1405 Asn Ile Pro Ile Ser Met Leu Asn Val Thr His Glu Thr Cys Val Pro 1410 1415 1420 Gly Tyr Trp Ile Ala Pro Val Gly Leu Thr Val Phe Met Leu Ala Thr 1430 1435 1425 Asn Val Leu Leu Met Asn Val Met Val Ala Gly Cys Thr Tyr Ile Phe 1445 1450 1455 Glu Lys His Ile Gln Ser Thr Arg Glu Ile Phe Leu Phe Glu Arg Tyr 1460 1465 1470 Gly Gln Val Met Glu Tyr Glu Ser Thr Pro Trp Leu Pro Pro Pro Phe 1480 1475 1485 Thr Ile Ile Tyr His Val Ile Trp Leu Phe Lys Leu Ile Lys Ser Ser 1495 1500 1490 Ser Arg Met Phe Glu Arg Lys Asn Leu Phe Asp Gln Ser Leu Lys Leu 1510 1515 Phe Leu Ser Pro Asp Glu Met Glu Lys Val His Thr Phe Glu Glu Glu 1525 1530 1535 Ser Val Glu Asp Met Lys Arg Glu Thr Glu Lys Lys Asn Leu Ser Ser 1540 1545 1550 Asn Asp Glu Arg Ile His Arg Thr Ala Glu Arg Thr Asp Ala Ile Leu 1555 1560 1565 Asn Arg Val Ser His Leu Thr Gln Leu Glu Phe Thr Leu Lys Glu Glu 1570 1575 1580 Ile Arg Glu Leu Glu His Lys Met Lys Asn Met Asp Ser Arg His Lys 1585 1590 1595 Glu Gln Met Asn Leu Met Leu Asp Met Asn Lys Lys Leu Gly Lys Phe 1605 1610 1615 Ile Ser Gly Lys Tyr Lys Arg Gly Ser Phe Gly Gly Ser Gly Ser Asp 1620 1625 1630 Gly Gly Gly Gly Ser Ser Asp Asn Ser Lys Leu Glu Pro Asn Asn Ser 1635 1640 1645 Val Pro Met Ile Thr Val Asp Gly Pro Ser Pro Ile Gly Ser Arg Arg 1655 1660 Thr Ser Gly Gln Tyr Leu Lys Arg Asp Ser Leu Gln Ala Lys Lys 1670 1675 Ile Thr Glu Asn Arg Arg Ser Ser Leu Glu Gln Pro Lys Ile Pro Ser 1685 1690 Ile Gln Phe Asn Leu Met Glu Asp Gln Asp Glu Ser Ala Ala Glu Ser 1700 1705 1710 Ala Thr Glu Glu Val Ser Ile Ser Ile Pro Val Pro Gln Met Arg Val 1715 1720 . 1725 Arg Gln Val Thr Glu Ser Asp Lys Ser Asp Leu Ser Glu Asp Asp Leu 1730 1735 1740 Ile Thr Arg Glu Asp Ala Pro Pro Thr Ser Ile Asn Leu Pro Arg Gly 1750 1755 Pro Arg Arg His Ala Leu Tyr Ser Thr Ile Ala Asp Ala Ile Glu Thr

-22-

1765 1770 1775

Glu Asp Asp Phe Tyr Ala Asp Ser Pro Val Pro Met Pro Met Thr Pro
1780 1785 1790

Val Gln Pro Ala Asp Gly Ser Phe Phe Gly Glu Asn Asp Ser Arg Tyr
1795 1800 1805

Gln Arg Asp Asp Ser Asp Tyr Glu
1810 1815

<210> 14 <211> 1387 <212> PRT <213> C. Elegans

<400> 14 Met Arg Lys Ser Arg Arg Val Arg Lys Leu Val Arg His Ala Ser Leu 10 Ile Glu Asn Ile Arg His Arg Thr Ser Ser Phe Leu Arg Leu Leu Asn 20 25 Ala Pro Arg Asn Ser Met Cys Asn Ala Asn Thr Val His Ser Ile Ser 35 40 45 Ser Phe Arg Ser Asp His Leu Ser Arg Lys Ser Thr His Lys Phe Leu 55 Asp Asn Pro Asn Leu Phe Ala Ile Glu Leu Thr Glu Lys Leu Ser Pro 70 75 Pro Trp Ile Glu Asn Thr Phe Glu Lys Arg Glu Cys Ile Arg Phe Ala 90 85 Ala Leu Pro Lys Asp Pro Glu Arg Cys Gly Cys Gly Arg Pro Leu Ser 100 105 Ala His Thr Pro Ala Ser Thr Phe Phe Ser Thr Leu Pro Val His Leu 120 125 Leu Glu Lys Glu Gln Gln Thr Trp Thr Ile Ala Asn Asn Thr Gln Thr 135 140 130 Ser Thr Thr Asp Ala Phe Gly Thr Ile Val Phe Gln Gly Gly Ala His 150 155 Ala His Lys Ala Gln Tyr Val Arg Leu Ser Tyr Asp Ser Glu Pro Leu 170 175 165 Asp Val Met Tyr Leu Met Glu Lys Val Trp Gly Leu Glu Ala Pro Arg 180 185 Leu Val Ile Thr Val His Gly Gly Met Ser Asn Phe Glu Leu Glu Glu 195 200 Arg Leu Gly Arg Leu Phe Arg Lys Gly Met Leu Lys Ala Ala Gln Thr 215 220 Thr Gly Ala Trp Ile Ile Thr Ser Gly Leu Asp Ser Gly Val Val Arg 230 235 His Val Ala Lys Ala Leu Asp Glu Ala Gly Ile Ser Ala Arg Met Arg 245 250 Ser Gln Ile Val Thr Ile Gly Ile Ala Pro Trp Gly Val Ile Lys Arg 260 265 270 Lys Glu Arg Leu Ile Arg Gln Asn Glu His Val Tyr Tyr Asp Val His 280 285 Ser Leu Ser Val Asn Ala Asn Val Gly Ile Leu Asn Asp Arg His Ser 295 300 Tyr Phe Leu Leu Ala Asp Asn Gly Thr Val Gly Arg Phe Gly Ala Asp 310 315 Leu His Leu Arg Gln Asn Leu Glu Asn His Ile Ala Thr Phe Gly Cys 330 325 335 Asn Gly Arg Lys Val Pro Val Val Cys Thr Leu Leu Glu Gly Gly Ile 340 345 350 Ser Ser Ile Asn Ala Ile His Asp Tyr Val Thr Met Lys Pro Asp Ile 360

-23-

									-23-						
Pro	Ala 370	Ile	Val	Cys	Asp	Gly 375	Ser	Gly	Arg	Ala	Ala 380	Asp	Ile	Ile	Ser
Phe 385	Ala	Ala	Arg	Tyr	Ile 390	Asn	Ser	Asp	Gly	Thr 395	Phe	Ala	Ala	Glu	Val 400
Gly	Glu	Lys	Leu	Arg 405	Asn	Leu	Ile	Lys	Met 410	Val	Phe	Pro	Glu	Thr 415	Asp
Gln	Glu	Glu	Met 420	Phe	Arg	Lys	Ile	Thr 425	Glu	Cys	Val	Ile	Arg 430	Asp	Asp
Leu	Leu	Arg 435	Ile	Phe	Arg	Tyr	Gly 440	Gln	Glu	Glu	Glu	Glu 445	Asp	Val	Asp
Phe	Val 450	Ile	Leu	Ser	Thr	Val 455	Leu	Gln	Lys	Gln	Asn 460	Leu	Pro	Pro	Asp
Glu 465	Gln	Leu	Ala	Leu	Thr 470	Leu	Ser	Trp	Asn	Arg 475	Val	Asp	Leu	Ala	Lys 480
Ser	Cys	Leu	Phe	Ser 485	Asn	Gly	Arg	Lys	Trp 490	Ser	Ser	Asp	Val	Leu 495	Glu
Lys	Ala	Met	Asn 500	Asp	Ala	Leu	Tyr	Trp 505	Asp	Arg	Val	Asp	Phe 510	Val	Glu
Cys	Leu	Leu 515	Glu	Asn	Gly	Val	Ser 520	Met	Lys	Asn	Phe	Leu 525	Ser	Ile	Asn
Arg	Leu 530	Glu	Asn	Leu	Tyr	Asn 535	Met	Asp	Asp	Ile	Asn 540	Ser	Ala	His	Ser
Val 545	Arg	Asn	Trp	Met	Glu 550	Asn	Phe	Asp	Ser	Met 555	Asp	Pro	His	Thr	Tyr 560
Leu	Thr	Ile	Pro	Met 565	Ile	Gly	Gln	Val	Val 570	Glu	Lys	Leu	Met	Gly 575	Asn
Ala	Phe	Gln	Leu 580	Tyr	Tyr	Thr	Ser	Arg 585	Ser	Phe	Lys	Gly	Lys 590	Tyr	Asp
_	_	595					600		_			Arg 605	_	_	_
	610		_			615	-	-	-		620	Asp			
625					630					635		Phe		_	640
		_		645				-	650	_		Ala		655	
_			660		_			665	_	-		Leu	670		_
		675					680			_		Tyr 685		_	
	690		_	_		695	_				700	Cys	_		_
705					710		_	-	_	715	_	His	_		720
	_			725					730	-	_	Asn		735	_
Leu	Ser	Leu	Ala 740	Val	Leu	Ala	Asn	Thr 745	Lys	Thr	Phe	Leu	Ala 750	His	Pro
Cys	Cys	Gln 755	Ile	Leu	Leu	Ala	Glu 760	Leu	Trp	His	Gly	Ser 765	Leu	Lys	Val
Arg	Ser 770	Gly	Ser	Asn	Val	Arg 775	Val	Leu	Thr	Ala	Leu 780	Ile	Cys	Pro	Pro
Ala 785	Ile	Leu	Phe	Met	Ala 790	Tyr	Lys	Pro	Lys	His 795	Ser	Lys	Thr	Ala	Arg 800
Leu	Leu	Ser	Glu		Thr	Pro	Glu	Gln	Leu 810	Pro	Tyr	Pro	Arg	Glu 815	Ser
				805					010					913	
Ile	Thr	Ser	Thr 820		Ser	Asn	Arg	Tyr 825		Tyr	Ser	Lys	Gly 830		Glu

-24-

Thr	Ile	Ile	Ser	Ser	Arg		Asn	Ser	Gly	Val		Ser	Val	Tyr	Gly
C	850 Ala	502	802	Mot	Mot	855 Pho	Tue	Ara	Glu	Pro	860 Gln	Lou	Acn	Luc	Pha
865					870		_	_		875				_	880
Glu	Arg	Phe	Arg	Ala 885	Phe	Tyr	Ser	Ser	Pro 890	Ile	Thr	Lys	Phe	Trp 895	Ser
Trp	Cys	Ile	Ala 900		Leu	Ile	Phe	Leu 905		Thr	Gln	Thr	Cys 910	Ile	Leu
Leu	Leu	Glu 915	Thr	Ser	Leu	Lys	Pro 920	Ser	Lys	Tyr	Glu	Trp 925	Ile	Thr	Phe
Ile	Tyr 930		Val	Thr	Leu	Ser 935	Val	Glu	His	Ile	Arg 940	Lys	Leu	Met	Thr
Ser 945	Glu	Gly	Ser	Arg	Ile 950	Asn	Glu	Lys	Val	Lys 955	Val	Phe	Tyr	Ala	Lys 960
Trp	Tyr	Asn	Ile	Trp 965	Thr	Ser	Ala	Ala	Leu 970	Leu	Phe	Phe	Leu	Val 975	Gly
Tyr	Gly	Phe	Arg 980	Leu	Val	Pro	Met	Tyr 985	Arg	His	Ser	Trp	Gly 990	Arg	Val
Leu	Leu	Ser 995	Phe	Ser	Asn	Val	Leu 1000		Tyr	Met	Lys	Ile 1005		Glu	Tyr
Leu	Ser 1010	Val	His	Pro	Leu	Leu 1015	Gly		Tyr	Ile	Gln 1020		Ala	Ala	Lys
Met	Val		Ser	Met			Ile	Cys	Val			Leu	Val	Pro	
1025	5 Ala	Dho	C1.,	Wa I	1030		Gl n	Λla	Lou	1035		Pro	Λen	Val	104
Met	нта	Pne	GIY	1045		Arg	GIII	нта	1050		GIU	FIO	ASII	1055	
Asp	Trp	His	Trp 1060		Leu	Val	Arg	Asn 1065	Ile		Tyr	Lys	Pro 1070		Phe
Met	Leu	Tyr 1075		Glu	Val	Tyr	Ala 1080		Glu	Ile	Asp	Thr 1085		Gly	Asp
Glu	Gly 1090		Arg	Cys	Phe	Pro 1095		Tyr	Phe	Ile	Pro 1100		Leu	Leu	Met
	Ile	Phe	Leu	Leu			Asn	Ile	Leu			Asn	Leu	Leu	
110		Dho	700	7.00	1110		7:00	7.00	802	1115		Tuc	Sor	Tue	112
AIA	Ile	Pne	ASII	1125		ıyı	MSII	Asp	1130		Giu	БУЗ	Ser	1135	
Ile	Trp	Leu	Phe	Gln	Arg	Tyr	Gln		Leu		Glu	Tyr			Ser
D	Phe	*	1140		D	Dha	C	1145		71.	uio	17-1	1150		Dho
		115	5				1160)				1165	5		
Ile	Asp 1170		Leu	Tyr	Asn	Leu 1175		Arg	Pro	Asp	Thr 1180		Arg	Pne	Arg
Ser 118	Glu		Ser	Ile	Lys 1190	Leu		Val	Thr	Glu 1195	Asp		Met	Lys	Arg 120
	Gln	Asp	Phe	Glu			Cys	Ile	Asp			Thr	Arg	Ile	
	Leu			1205	5				1210)				1215	5
	neu	гуу			TIII	БАЗ	GIU		neu	JCI	VOI	1111			1111
Clu			1220					1225	5				1230		
	Leu	123	Cys 5	Gln			1240	Asp	Leu			1245	Asn	Phe	
	Lys	1235 Ser	Cys 5	Gln			1240 Ile	Asp	Leu			1245 Asp	Asn	Phe	
Leu	Lys 1250 Ser	123! Ser)	Cys 5 Arg	Gln Val	Tyr	Asp 1255 Val	1240 Ile	Asp) Glu	Leu Thr	Lys	Ile 1260 Asn	1245 Asp)	Asn His	Phe Ile	Ser
Leu Asn 126	Lys 1250 Ser	123! Ser) Ser	Cys 5 Arg Asp	Gln Val Glu Ala	Tyr Val 1270 Ser	Asp 1255 Val	1240 Ile Gln	Asp) Glu Ile	Leu Thr Leu Leu	Lys Lys 1275 Pro	Ile 1260 Asn	1245 Asp) Lys	Asn His Lys	Phe Ile Leu Ile	Ser Ser 128 Glu
Leu Asn 126 Gln	Lys 1250 Ser 5	123! Ser) Ser Phe	Cys Arg Asp Ala Ile	Gln Val Glu Ala 1285 Thr	Tyr Val 1270 Ser	Asp 1255 Val) Ser	1240 Ile Gln Leu	Asp Glu Ile Ser	Leu Thr Leu Leu 1290 Asp	Lys Lys 1275 Pro	Ile 1260 Asn Asp	1245 Asp) Lys Thr	Asn His Lys Ser	Phe Ile Leu Ile 1295 Pro	Ser Ser 128 Glu
Leu Asn 126 Gln Val	Lys 1250 Ser 5 Asn	123! Ser Ser Phe	Cys Arg Asp Ala Ile 1300	Gln Val Glu Ala 1285 Thr	Tyr Val 1270 Ser Lys	Asp 1255 Val) Ser Thr	1240 Ile Gln Leu	Asp Glu Ile Ser Ile 130	Leu Thr Leu Leu 1290 Asp	Lys Lys 1275 Pro) Cys	Ile 1260 Asn Asp His	1245 Asp) Lys Thr	Asn His Lys Ser Ser 1310	Phe Ile Leu Ile 1295 Pro	Ser Ser 128 Glu Val

-25-

Gln Arg Asp His Thr Leu Arg Lys Leu Pro Thr Trp Glu Thr Ser Thr 1330

Ala Ser Thr Ser Ser Phe Glu Phe Val Phe Tyr Phe Thr Arg His Glu 1345

Gly Asn Glu Asn Lys Tyr Glu Phe Lys Leu Glu Lys Gly Gly Phe 1365

Trp Arg Asn Asn Tyr Val Ile Ser Trp Arg Leu 1380

<210> 15 <211> 1868 <212> PRT <213> C. Elegans

<400> 15 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$ Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln 20 25 Ser Lys Gly Gly Asp Gln Val Pro Pro Thr Ser Thr Thr Thr Gly Gly 35 40 Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala 55 Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp 75 70 Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val 90 85 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met 105 110 Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu 120 125 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys 135 140 Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu 150 155 Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser 165 170 175 Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys 185 190 Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu 195 200 205 Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro 215 220 Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr 230 235 240 Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu 245 250 255 Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys 260 265 270 Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr 275 280 285 280 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His 295 300 Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile 315 310 Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Leu Lys Lys Arg Glu Asp 330 325 335 Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Lys 340 345 Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

		355					360					365			
Asp	Asn 370	Gly	Thr	Val	Gly	Arg 375	Tyr	Gly	Ala	Glu	Val 380	Ile	Leu	Arg	Lys
Arg 385	Leu	Glu	Met	Tyr	Ile 390	Ser	Gln	Lys	Gln	Lys 395	Ile	Phe	Gly	Gly	Thr 400
Arg	Ser	Val	Pro	Val 405	Val	Cys	Val	Val	Leu 410	Glu	Gly	Gly	Ser	Cys 415	Thr
Ile	Arg	Ser	Val 420	Leu	Asp	Tyr	Val	Thr 425	Asn	Val	Pro	Arg	Val 430	Pro	Val
Val	Val	Cys 435	Asp	Gly	Ser	Gly	Arg 440	Ala	Ala	Asp	Leu	Leu 445	Ala	Phe	Ala
	450					455	_				460	_		Arg	-
465					470					475				Ala	480
				485				•	490					Lys 495	
Leu	Leu	Thr	11e 500	Phe	Arg	Leu	Gly	Glu 505	Gln	Gly	Glu	His	Asp 510	Val	Asp
		515					520	-	-			525		Ala	
	530					535					540			Ala	
545	_				550	_			_	555				Leu	560
				565					570	_		-		Val 575	_
			580		_			585		_			590	Ile	
		595					600					605		Thr	
	610			_	_	615		_		_	620	_	_	Arg	
625			_		630					635			_	Asn -	640
				645					650					Lys 655	
_		_	660		_			665	-	-			670	Val	
		675					680		_			685	_	Gln	
	690	_		_	_	695	_		_		700			Gly	
705	_				710			_		715		-		His	720
	•			725		_		-	730					11e 735	
_	_		740		_	_		745					750	Arg	
		755					760					765		Glu	
	770					775	_	_			780			Asp	
785	_		_		790				_	795				Lys	800
				805	_				810	_				Met 815	
			820					825					830	Glu -	
Ala	GIU	Asp	Tyr	Leu	GLu	val	GLu	ITe	Cys	GIU	GLu	Leu	Lys	Lys	Tyr

			840					845			
Ala Glu Glu Pl 850	he Arg 1	Ile Leu 855	Ser	Leu	Glu	Leu	Leu 860	Asp	His	Cys	Tyr
His Val Asp A: 865	-	Gln Thr 370	Leu	Gln	Leu	Leu 875	Thr	Tyr	Glu	Leu	Ser 880
Asn Trp Ser A	sn Glu 7 885	Thr Cys	Leu	Ala	Leu 890	Ala	Val	Ile	Val	Asn 895	Asn
Lys His Phe Lo	eu Ala H 00	His Pro	-	Cys 905	Gln	Ile	Leu	Leu	Ala 910	Asp	Leu
Trp His Gly G 915	ly Leu <i>P</i>	Arg Met	Arg 920	Thr	His	Ser	Asn	Ile 925	Lys	Val	Val
Leu Gly Leu I. 930	le Cys E	Pro Pro 935	Phe	Ile	Gln	Met	Leu 940	Glu	Phe	Lys	Thr
Arg Glu Glu Lo 945		Asn Gln 950	Pro	Gln	Thr	Ala 955	Ala	Glu	His	Gln	Asn 960
Asp Met Asn T	965			•	970					975	
_	80	-		985			_	_	990	-	
Asn Asn Ala H. 995		_	1000	1	_		_	1005	ò		
Gly Ser Ala G		1015	5				1020)		_	
Arg Lys Ala Ly	1	1030	_	_	_	1035	•		_		104
Ala Cys Glu A	1045				1050)				1055	•
	060	-		1065	5				1070)	
Tyr Met Arg A	ia Asn s	ser Arg	1080	_	Tyr	ASI	ASN	1085		Asp	Met
Ser Lys Thr So 1090		1095	5				1100)			
1090 Leu Gln Lys So 1105	er Asn I 1	1095 Lle Thr 1110	Ser	Thr	Asp	Arg 1115	1100 Pro) Asn	Pro	Met	Glu 112
1090 Leu Gln Lys Sc 1105 Gln Phe Gln G	er Asn I 1 ly Thr F 1125	1095 Lle Thr Ll10 Arg Lys	Ser Ile	Thr Lys	Asp Met 1130	Arg 1115 Arg	1100 Pro Arg	Asn Arg	Pro Phe.	Met Tyr 1135	Glu 112 Glu
1090 Leu Gln Lys Sci 1105 Gln Phe Gln G. Phe Tyr Ser A.	er Asn I 1 ly Thr A 1125 la Pro I 140	1095 Tle Thr 1110 Arg Lys	Ser Ile Thr	Thr Lys Phe 1145	Asp Met 1130 Trp	Arg 1115 Arg Ser	1100 Pro Arg Trp	Asn Arg Thr	Pro Phe	Met Tyr 1135 Ser	Glu 112 Glu Phe
1090 Leu Gln Lys So 1105 Gln Phe Gln G Phe Tyr Ser A 11 Ile Leu Phe I 1155	er Asn I 1 1y Thr A 1125 la Pro I 140 le Thr E	1095 The Thr 1110 Arg Lys The Ser	Ser Ile Thr Thr 1160	Thr Lys Phe 1145 Tyr	Asp Met 1130 Trp Thr	Arg 1115 Arg Ser Leu	1100 Pro Arg Trp Leu	Asn Arg Thr Val 1165	Pro Phe Ile 1150 Lys	Met Tyr 1135 Ser Thr	Glu 112 Glu Phe Pro
1090 Leu Gln Lys So 1105 Gln Phe Gln G: Phe Tyr Ser A: Ile Leu Phe I: 1155 Pro Arg Pro Ti 1170	er Asn I ly Thr F 1125 la Pro I 140 le Thr E	1095 The Thr 1110 Arg Lys The Ser The Phe The Glu 1175	Ser Ile Thr Thr 1160 Tyr	Thr Lys Phe 1145 Tyr	Asp Met 1130 Trp Thr	Arg 1115 Arg Ser Leu	1100 Pro Arg Trp Leu Ala 1180	Asn Arg Thr Val 1165 Tyr	Pro Phe Ile 1150 Lys Val	Met Tyr 1135 Ser Thr	Glu 112 Glu Phe Pro
1090 Leu Gln Lys So 1105 Gln Phe Gln G. Phe Tyr Ser A. Ile Leu Phe I. 1155 Pro Arg Pro T. 1170 Phe Gly Leu G. 1185	er Asn I ly Thr F 1125 la Pro I 140 le Thr E hr Val I lu Gln V	1095 Ile Thr 1110 Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg	Ser Ile Thr Thr 1160 Tyr Lys	Thr Lys Phe 1145 Tyr Ile Ile	Asp Met 1130 Trp Thr Leu Ile	Arg 1115 Arg Ser Leu Ile Met 1195	Arg Trp Leu Ala 1180 Ser	Asn Arg Thr Val 1165 Tyr Asp	Pro Phe Ile 1150 Lys Val Ala	Met Tyr 1135 Ser Thr Ala Lys	Glu 112 Glu Phe Pro Ala Pro 120
1090 Leu Gln Lys Se 1105 Gln Phe Gln G. Phe Tyr Ser A. 11155 Pro Arg Pro The 1170 Phe Gly Leu G. 1185 Phe Tyr Glu Ly	er Asn I I Ily Thr F I125 la Pro I I40 le Thr E hr Val I lu Gln V I ys Ile F 1205	1095 Ile Thr 1110 Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg 1190 Arg Thr	Ser Ile Thr Thr 1160 Tyr Lys	Thr Lys Phe 1145 Tyr Ile Ile Val	Asp Met 1130 Trp Thr Leu Ile Cys 1210	Arg 1115 Arg Ser Leu Ile Met 1195 Ser	Arg Trp Leu Ala 1180 Ser	Asn Arg Thr Val 1165 Tyr Asp	Pro Phe 1150 Lys Val Ala Asn	Met Tyr 1135 Ser Thr Ala Lys Cys 1215	Glu 112 Glu Phe Pro Ala Pro 120 Val
1090 Leu Gln Lys Sci 1105 Gln Phe Gln G. Phe Tyr Ser A. Ile Leu Phe I. 1155 Pro Arg Pro The 1170 Phe Gly Leu G. 1185 Phe Tyr Glu Ly Thr Ile Leu A.	er Asn I I Ily Thr F I125 la Pro I I40 le Thr E hr Val I lu Gln V I ys Ile F 1205	1095 Ile Thr 1110 Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg 1190 Arg Thr	Ser Ile Thr Thr 1160 Tyr Lys Tyr	Thr Lys Phe 1145 Tyr Ile Ile Val	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val	Arg 1115 Arg Ser Leu Ile Met 1195 Ser	Arg Trp Leu Ala 1180 Ser Phe	Asn Arg Thr Val 1165 Tyr Asp Trp	Pro Phe Ile 1150 Lys Val Ala Asn	Met Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg	Glu 112 Glu Phe Pro Ala Pro 120 Val
1090 Leu Gln Lys Sci 1105 Gln Phe Gln G. Phe Tyr Ser A. Ile Leu Phe I. 1155 Pro Arg Pro The 1170 Phe Gly Leu G. 1185 Phe Tyr Glu Ly Thr Ile Leu A.	er Asn I I Ily Thr F I125 la Pro I 140 le Thr E hr Val I lu Gln V I ys Ile F 1205 la Ile I	1095 Ile Thr 1110 Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg 1190 Arg Thr	Ser Ile Thr 1160 Tyr Lys Tyr	Thr Lys Phe 1145 Tyr Ile Ile Val Ile 1225 Val	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val	Arg 1115 Arg Ser Leu Ile Met 1195 Ser	Arg Trp Leu Ala 1180 Ser Phe	Asn Arg Thr Val 1165 Tyr Asp Trp Phe	Pro Phe 1150 Lys Val Ala Asn Met 1230 Asp	Met Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg	Glu 112 Glu Phe Pro Ala Pro 120 Val
1090 Leu Gln Lys So 1105 Gln Phe Gln G Phe Tyr Ser A 11 Ile Leu Phe I 1155 Pro Arg Pro Ti 1170 Phe Gly Leu G 1185 Phe Tyr Glu L Thr Ile Leu A 12 Phe Gly Ser V	er Asn I ly Thr F 1125 la Pro I 140 le Thr E hr Val I lu Gln V 1205 la Ile I 220 al Ala I	Ile Thr Illo Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg Il90 Arg Thr Ile Phe	Ser Ile Thr Thr 1160 Tyr Lys Tyr Tyr Arg 1240 Asp	Thr Lys Phe 1145 Tyr Ile Ile Val Ile 1225 Val	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val	Arg 1115 Arg Ser Leu Ile Met 1195 Ser Gly Leu	Arg Trp Leu Ala 1180 Ser Phe Ala	Asn Arg Thr Val 1165 Tyr Asp Trp Phe Cys 1245 His	Pro Phe 1150 Lys Val Ala Asn Met 1230 Asp	Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg	Glu 112 Glu Phe Pro Ala Pro 120 Val Cys
1090 Leu Gln Lys So 1105 Gln Phe Gln G: Phe Tyr Ser A: Ile Leu Phe I: 1155 Pro Arg Pro T: 1170 Phe Gly Leu G: 1185 Phe Tyr Glu Ly Thr Ile Leu A: Phe Gly Ser Va 1235 Leu Trp Thr Ma 1250 Gly Pro Tyr Va 1265	er Asn I ly Thr F 1125 la Pro I 140 hr Val I lu Gln V 1205 la Ile I 220 al Ala T et Lys I al Thr M	Ile Thr Illo Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg Il90 Arg Thr Ile Phe Tyr Gly Leu Leu 1255 Met Ala	Ser Ile Thr Thr 1160 Tyr Lys Tyr Tyr Arg 1240 Asp Gly	Thr Lys Phe 1145 Tyr Ile Val Ile 1225 Val Tyr Lys	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val Ile Met Met	Arg 1115 Arg Ser Leu Ile Met 1195 Ser Gly Leu Ser Ile 1275	Arg Trp Leu Ala 1180 Ser Phe Ala Val 1260 Gln	Asn Arg Thr Val 1165 Tyr Asp Trp Phe Cys 1245 His	Pro Phe 1150 Lys Val Ala Asn Met 1230 Asp Pro Met	Met Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg Ser Lys	Glu 112 Glu Phe Pro Ala Pro 120 Val Cys Val Leu Tyr 128
1090 Leu Gln Lys So 1105 Gln Phe Gln G: Phe Tyr Ser A: Ile Leu Phe I: 1155 Pro Arg Pro I: 1170 Phe Gly Leu G: 1185 Phe Tyr Glu L: Thr Ile Leu A: 1235 Leu Trp Thr Mo 1250 Gly Pro Tyr Va 1265 Ile Ile Val Mo	er Asn I ly Thr F 1125 la Pro I 140 le Thr E hr Val I lu Gln V 1205 la Ile I 220 al Ala I et Lys I al Thr M 1et Leu V 1285	Ile Thr Illo Arg Lys Ile Ser Phe Phe Ile Glu 1175 Val Arg Il90 Arg Thr Ile Phe Tyr Gly Leu Leu 1255 Met Ala I270 Val Val	Ser Ile Thr Thr 1160 Tyr Lys Tyr Arg 1240 Asp Gly Thr	Thr Lys Phe 1145 Tyr Ile Val Ile 1225 Val Tyr Lys Leu	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val Ile Met Met Leu 1290	Arg 1115 Arg Ser Leu Ile Met 1195 Ser Gly Leu Ser Ile 1275	Arg Trp Leu Ala 1180 Ser Phe Ala Val 1260 Gln Phe	Asn Arg Thr Val 1165 Tyr Asp Trp Phe Cys 1245 His Asn Gly	Pro Phe 1150 Lys Val Ala Asn Met 1230 Asp Pro Met Leu	Met Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg Ser Lys Ser Ala 1295	Glu 112 Glu Phe Pro Ala Pro 120 Val Cys Val Leu Tyr 128 Arg
1090 Leu Gln Lys So 1105 Gln Phe Gln G. Phe Tyr Ser A. Ile Leu Phe II 1155 Pro Arg Pro II 1170 Phe Gly Leu G. 1185 Phe Tyr Glu Ly Thr Ile Leu A. 1235 Leu Trp Thr Mo 1250 Gly Pro Tyr Va 1265 Ile Ile Val Mo Gln Ser Ile II	er Asn I 1y Thr A 1125 la Pro I 140 le Thr E hr Val I lu Gln V 1205 la Ile I 220 al Ala I et Lys I et Leu V 1285 hr Tyr E	Ile Thr Illo Arg Lys Ile Ser Phe Phe Ile Glu 1175 Al Arg Illo Arg Thr Ile Phe Tyr Gly Leu 1255 Met Ala 1270 Val Val	Thr 1160 Tyr Lys Tyr Arg 1240 Asp Gly Thr	Thr Lys Phe 1145 Tyr Ile Val Ile 1225 Val Tyr Lys Leu Thr 1305	Asp Met 1130 Trp Thr Leu Ile Cys 1210 Val Ile Met Met Leu 1290 Trp	Arg 1115 Arg Ser Leu Ile Met 1195 Ser Gly Leu Ser Ile 1275 Ser	Arg Trp Leu Ala 1180 Ser Phe Ala Val 1260 Gln Phe	Asn Arg Thr Val 1165 Tyr Asp Trp Phe Cys 1245 His Asn Gly Ile	Pro Phe 1150 Lys Val Ala Asn Met 1230 Asp Pro Met Leu 1310	Met Tyr 1135 Ser Thr Ala Lys Cys 1215 Arg Ser Lys Ser Ala 1295 Val	Glu 112 Glu Phe Pro Ala Pro 120 Val Cys Val Leu Tyr 128 Arg

1320 1325 Asp Glu Ile Asp Thr Cys Gly Asp Glu Ala Trp Asp Gln His Leu Glu 1335 1340 Asn Gly Gly Pro Val Ile Leu Gly Asn Gly Thr Thr Gly Leu Ser Cys 1350 1355 Val Pro Gly Tyr Trp Ile Pro Pro Leu Leu Met Thr Phe Phe Leu Leu 1365 1370 Ile Ala Asn Ile Leu Leu Met Ser Met Leu Ile Ala Ile Phe Asn His 1380 1385 1390 Ile Phe Asp Ala Thr Asp Glu Met Ser Gln Gln Ile Trp Leu Phe Gln 1400 1395 1405 Arg Tyr Lys Gln Val Met Glu Tyr Glu Ser Thr Pro Phe Leu Pro Pro 1410 1415 1420 Pro Leu Thr Pro Leu Tyr His Gly Val Leu Ile Leu Gln Phe Val Arg 1425 1430 1435 Thr Arg Leu Ser Cys Ser Lys Ser Gln Glu Arg Asn Pro Ile Leu Leu 1445 1450 Leu Lys Ile Ala Glu Leu Phe Leu Asp Asn Asp Gln Ile Glu Lys Leu 1465 1460 1470 His Asp Phe Glu Glu Asp Cys Met Glu Asp Leu Ala Arg Gln Lys Leu 1475 1480 1485 Asn Glu Lys Asn Thr Ser Asn Glu Gln Arg Ile Leu Arg Ala Asp Ile 1490 1495 1500 Arg Thr Asp Gln Ile Leu Asn Arg Leu Ile Asp Leu Gln Ala Lys Glu 1510 1515 Ser Met Gly Arg Asp Val Ile Asn Asp Val Glu Ser Arg Leu Ala Ser 1525 1530 1535 Val Glu Lys Ala Gln Asn Glu Ile Leu Glu Cys Val Arg Ala Leu Leu 1540 1545 1550 Asn Gln Asn Asn Ala Pro Thr Ala Ile Gly Arg Cys Phe Ser Pro Ser 1560 1555 1565 Pro Asp Pro Leu Val Glu Thr Ala Asn Gly Thr Pro Gly Pro Leu Leu 1570 1575 1580 Leu Lys Leu Pro Gly Thr Asp Pro Ile Leu Glu Glu Lys Asp His Asp 1590 1595 Ser Gly Glu Asn Ser Asn Ser Leu Pro Pro Gly Arg Ile Arg Asn 1605 1610 Arg Thr Ala Thr Ile Cys Gly. Gly Tyr Val Ser Glu Glu Arg Asn Met 1620 1625 1630 Met Leu Leu Ser Pro Lys Pro Ser Asp Val Ser Gly Ile Pro Gln Gln 1635 1640 1645 Arg Leu Met Ser Val Thr Ser Met Asp Pro Leu Pro Leu Pro Leu Ala 1655 1660 Lys Leu Ser Thr Met Ser Ile Arg Arg Arg His Glu Glu Tyr Thr Ser 1670 1675 Ile Thr Asp Ser Ile Ala Ile Arg His Pro Glu Arg Arg Ile Arg Asn 1685 1690 1695 Asn Arg Ser Asn Ser Ser Glu His Asp Glu Ser Ala Val Asp Ser Glu 1700 1705 1710 Gly Gly Asn Val Thr Ser Ser Pro Arg Lys Arg Ser Thr Arg Asp 1715 1720 1725 Leu Arg Met Thr Pro Ser Ser Gln Val Glu Glu Ser Thr Ser Arg Asp 1730 1735 1740 Gln Ile Phe Glu Ile Asp His Pro Glu His Glu Glu Asp Glu Ala Gln 1750 1755 Ala Asp Cys Glu Leu Thr Asp Val Ile Thr Glu Glu Glu Asp Glu Glu 1770 1765 1775 Glu Asp Asp Glu Glu Asp Asp Ser His Glu Arg His His Ile His Pro 1790 1780 1785 Arg Arg Lys Ser Ser Arg Gln Asn Arg Gln Pro Ser His Thr Leu Glu

<221> unsure

```
-29-
                          1800
                                              1805
Thr Asp Leu Ser Glu Gly Glu Glu Val Asp Pro Leu Asp Val Leu Lys
                     1815
                                1820
Met Lys Glu Leu Pro Ile Ile His Gln Ile Leu Asn Glu Glu Gln
                   1830
                                      1835
Ala Gly Ala Pro His Ser Thr Pro Val Ile Ala Ser Pro Ser Ser Ser
              1845
                                1850
Arg Ala Asp Leu Thr Ser Gln Lys Cys Ser Asp Val
                               1865
     <210> 16
     <211> 489
     <212> DNA
    <213> Mus Musculus
     <400> 16
ccctgaaaga ctcgacttct gctgctagcg ctggagctga gttagttttg agaaggtttc
ceggggctgt cettgttegg tggcccgtgc cacegcetec ggagacgett teegatagat
                                                                    120
qqctqcaqqc cqcqqaggtg gaggaggagc cgctqccctt ccqqagtccg ccccqtgagg
                                                                    180
                                                                    240
agaatgtccc agaaatcctg gatagagagc actttgacca agagggagtg tgtatatatt
ataccaaget ccaaagacce tcacagatgt cttccaggat gtcagatttg tcagcaactt
                                                                    300
                                                                    360
gtcagatgtt tctgtggtcg tttggtcaag caacatgcat gctttactgc aagtcttgcc
atgaaatact cagatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct
                                                                    420
                                                                    480
gtggaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg
                                                                    489
ggttctcat
     <210> 17
     <211> 102
     <212> PRT
     <213> Mus Musculus
     <400> 17
Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys
                5
                                  10
Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly
                                                  30
        20
                              25
Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val
      35
                         40
                                             ·45
Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp
                      55
   50
                                          60
Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val
                 70
                                    75
Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn
              85
                                   90
Phe Gln Gly Gly Ser His
           100
      <210> 18
      <211> 410
      <212> DNA
      <213> Homo Sapiens
     <220>
      <221> unsure
      <222> (6)...(6)
      <221> unsure
      <222> (58)...(58)
```

-30-

```
<222> (89)...(89)
       <221> unsure
      <222> (406)...(406)
      <400> 18
gccgcnggag cctgagcgga gggtgtgcgc agcctcgcca gcgggggccc cgggctgngc
                                                                          60
cattgeetea etgageeage geetgeetne taeetegeeg acagetggaa ecagtgegae
                                                                        120
ctagtggctc tcacctgctt cctcctgggc gtgggctgcc ggctgacccc gggtttgtac
                                                                         180
cacctgggcc gcactgtcct ctgcatcgac ttcatggttt tcacggtgcg gctgcttcac
                                                                        240
atcttcacgg tcaacaaca gctggggccc aagatcgtca tcgtgagcaa gatgatgaag
                                                                         300
gacgtgttct tcttcctctt cttcctcggc gtgtggctgg tagctatggg ttgggccacg
                                                                        360
gaggggttcc tgaggccacg ggacagtgac ttcccaagta tcctgncgcc
                                                                         410
       <210> 19
       <211> 131
       <212> PRT
       <213> Homo Sapiens
      <220>
       <221> UNSURE
      <222> (15)...(15)
<223> UNKNOWN
      <221> UNSURE
      <222> (25)...(25)
<223> UNKNOWN
      <221> UNSURE
      <222> (131)...(131)
      <223> UNKNOWN
      <400> 19
Ala Glu Gly Val Arg Ser Leu Ala Ser Gly Gly Pro Gly Leu Xaa His
. 1
                  5
                                     10
                                                          15
Cys Leu Thr Glu Pro Ala Pro Ala Xaa Tyr Leu Ala Asp Ser Trp Asn
            20
                                 25
                                                      30
Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys
                             40
                                                  45
Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile
                        55
                                             60
Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn
                     70
                                         75
Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp
                 85
                                     90
Val Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Met Gly
            100
                                 105
Trp Ala Thr Glu Gly Phe Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser
        115
Ile Leu Xaa
    130
      <210> 20
      <211> 389
      <212> DNA
      <213> Homo Sapiens
      <400> 20
                                                                         60
caaatttttt gttagtacac catctcatcc aaattgcaaa agtcacatgg aaactggaac
caaagatcaa gaaactgttt gctctaaagc tacagaagga gataatacag aatttggagc
                                                                        120
```

-31-

```
180
atttgtagga cacagagata gcatggattt acagaggttt aaagaaacat caaacaagat
                                                                    240
aaaaatacta tccaataaca atacttctga aaacactttg aaacgagtga gttctcttgc
                                                                    300
tggatttact gactgtcaca gaacttccat tcctgttcat tcaaaacgag aaaagatcag
                                                                    360
tagaaggcca tctaccgaag acactcatga agtagattcc aaagcagctt taataccggt
ttgtagattt caactaaaca gatatatat
                                                                    389
      <210> 21
      <211> 415
      <212> DNA
      <213> Homo Sapiens
      <400> 21
atttctagtt tttcaaattt gccagtcttt ttgaatagta tctccttctt ttctcatgtt
                                                                     60
ttatatttaa aacttttta tgtccatcat cactttaaac atacttattt tgtcatctat
                                                                    120
aaccaataat tccactatct tatcagaaat caaataccgt ttatgtaagt tgactcccat
                                                                    180
gagttctaaa ttgccattgt gaggtcatct tcggttaggc tttaatttgt tgcaaagttg
                                                                    240
tgcagctcag ggtcaggaag agtccctcca gaaaggagga tttgttactg tgaatctctt
                                                                    300
tgttaactaa cctctttccc cactgaaata acttttttca ataacatgat tttaacaaca
                                                                    360
taatctctct atgccagaac agatatatat gaatgtaagt caatattttc ttgag
                                                                    415
      <210> 22
      <211> 405
      <212> DNA
      <213> Mus Musculus
      <400> 22
ttattatggc ttatcatgaa aaaccagtcc tgcctcctcc tcttatcatc ctcagccata
                                                                     60
tagtttcact gttttgctgt gtatgcaaaa gaagaaagaa agataagact tccgatgggc
                                                                    120
caaaactttt cttaacagaa gaagatcaaa agaaactcca tgattttgaa gagcagtgtg
                                                                    180
                                                                    240
300
gggtcacttt tgaaagagtg gagcagatga gcattcagat taaagaagtt ggagatcgtg
tcaactacat aaaaagatca ttacagtctt tagattctca aattggtcat ctgcaagatc
                                                                    360
tctcagccct aacagtagat acattgaaaa cacttacagc ccaga
                                                                    405
      <210> 23
      <211> 5117
      <212> DNA
      <213> Homo Sapiens
      <220>
      <221> unsure
      <222> (2382)...(2382)
      <223> unknown
     <221> unsure
      <222> (4664)...(4664)
     <223> unknown
      <221> unsure
      <222> (4682)...(4682)
      <223> unknown
     <221> unsure
      <222> (4702)...(4702)
     <223> unknown
      <221> unsure
      <222> (5038)...(5039)
      <223> unknown
```

-32-

<221> unsure <222> (5056)...(5056) <223> unknown

<221> unsure

<222> (5071)...(5072)

<400> 23

gatggcaaca tggtgaagaa tcaatggcta aagcattagt tgcctgtaag atctatcgtt 60 caatggcata tgaagcaaag cagagtgacc tggtagatga tacttcagaa gaactaaaac 120 agtattccaa tgattttggt cagttggccg ttgaattatt agaacagtcc ttcagacaag 180 atgaaaccat ggctatgaaa ttgctcactt atgaactgaa gaactggagt aattcaacct 240 gccttaagtt agcagtttct tcaagactta gaccttttgt agctcacacc tgtacacaaa 300 tgttgttatc tgatatgtgg atgggaaggc tgaatatgag gaaaaattcc tggtacaagg 360 tcatactaag cattttagtt ccacctgcca tattgctgtt agagtataaa actaaggctg 420 aaatgtccca tatcccacaa tctcaagatg ctcatcagat gacaatggat gacagcgaaa 480 acaactttca gaacataaca gaagagatcc ccatggaagt gtttaaagaa gtacggattt 540 tggatagtaa tgaaggaaag aatgagatgg agatacaaat gaaatcaaaa aagcttccaa 600 ttacgcgaaa gttttatgcc ttttatcatg caccaattgt aaaattctgg tttaacacgt 660 tggcatattt aggatttctg atgctttata catttgtggt tcttgtacaa atggaacagt 720 taccttcagt tcaagaatgg attgttattg cttatatttt tacttatgcc attgagaaag 780 tccgtgagat ctttatgtct gaagctggga aagtaaacca gaagattaaa gtatggttta 840 gtgattactt caacatcagt gatacaattg ccataatttc tttcttcatt ggatttggac 900 taagatttgg agcaaaatgg aactttgcaa atgcatatga taatcatgtt tttgtggctg 960 gaagattaat ttactgtctt aacataatat tttggtatgt gcgtttgcta gattttctag 1020 ctgtaaatca acaggcagga ccttatgtaa tgatgattgg aaaaatggtg gccaatatgt 1080 tctacattgt agtgattatg gctcttgtat tacttagttt tggtgttccc agaaaggcaa 1140 tactttatcc tcatgaagca ccatcttgga ctcttgctaa agatatagtt tttcacccat 1200 actggatgat ttttggtgaa gtttatgcat acgaaattga tgtgtgtgca aatgattctg 1260 ttatecetea aatetgtggt eetgggaegt ggttgaetee atttetteaa geagtetaee 1320 tctttgtaca gtatatcatt atggttaatc ttcttattgc atttttcaac aatgtgtatt 1380 tacaagtgaa ggcaatttcc aatattgtat ggaagtacca gcgttatcat tttattatgg 1440 cttatcatga gaaaccagtt ctgcctcctc cacttatcat tcttagccat atagtttctc 1500 tgttttgctg catatgtaag agaagaaaga aagataagac ttccgatgga ccaaaacttt 1560 tcttaacaga agaagatcaa aagaaacttc atgattttga agagcagtgt gttgaaatgt 1620 atttcaatga aaaagatgac aaatttcatt ctgggagtga agagagaatt cgtgtcactt 1680 ttgaaagagt ggaacagatg tgcattcaga ttaaagaagt tggagatcgt gtcaactaca 1740 taaaaagatc attacaatca ttagattctc aaattggcca tttgcaagat ctttcagccc 1800 tgacggtaga tacattaaaa acactcactg cccagaaagc gtcggaagct agcaaagttc 1860 ataatgaaat cacacgagaa ctgagcattt ccaaacactt ggctcaaaac cttattgatg 1920 atggtcctgt aagaccttct gtatggaaaa agcatggtgt tgtaaataca cttagctcct 1980 ctettectea aggtgatett gaaagtaata ateettttea ttgtaatatt ttaatgaaag 2040 atgacaaaga tccccagtgt aatatatttg gtcaagactt acctgcagta ccccagagaa 2100 aagaatttaa ttttccagag gctggttcct cttctggtgc cttattccca agtgctgttt 2160 cccctccaga actgcgacag agactacatg gggtagaact cttaaaaaata tttaataaaa 2220 atcaaaaatt aggcagttca tctactagca taccacatct gtcatcccca ccaaccaaat 2280 tttttgttag tacaccatct cagccaagtt gcaaaagcca cttggaaact ggaaccaaag 2340 atcaagaaac tgtttgctct aaagctacag aaggagataa tncagaattt ggagcatttg 2400 taggacacag agatagcatg gatttacaga ggtttaaaga aacatcaaac aagataaaaa 2460 tactatccaa taacaatact totgaaaaca otttgaaacg agtgagttot ottgotggat 2520 ttactgactg tcacagaact tccattcctg ttcattcaaa acaagcagaa aaaatcagta 2580 2640 gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta ataccggatt ggttacaaga tagaccatca aacagagaaa tgccatctga agaaggaaca ttaaatggtc 2700 tcacttctcc atttaagcca gctatggata caaattacta ttattcagct gtggaaagaa 2760 ataacttgat gaggttatca cagagcattc catttacacc tgtgcctcca agaggggagc 2820 2880 ctgtcacagt gtatcgtttg gaagagagtt cacccaacat actaaataac agcatgtctt cttggtcaca actaggcctc tgtgccaaaa tagagttttt aagcaaagag gagatgggag 2940 3000 gaggtttacg aagagctgtc aaagtacagt gtacctggtc agaacatgat atcctcaaat cagggcatct ttatattatc aaatcttttc ttccagaggt ggttaataca tggtcaagta 3060 tttataaaga agatacagtt ctgcatctct gtctgagaga aattcaacaa cagagagcag 3120

```
3180
cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa
ggttccttga agttttcctg ctgtattgcc attcagcagg acagtggttt gctgtggaag
                                                                      3240
                                                                      3300
aatgtatgac tggagaattt agaaaataca acaataataa tggagatgag attattccaa
ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag
                                                                      3360
                                                                      3420
gggagttact ggtacttgat ttgcaaggtg ttggtgaaaa tttgactgac ccatctgtga
taaaagcaga agaaaagaga teetgtgata tggtttttgg eecagcaaat etaggagaag
                                                                      3480
atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac
                                                                      3540
ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttcct caggatgagc
                                                                      3600
cttcagattt gaatcttcag cctggaaatt ccaccaaaga atcagaatca gctaattctg
                                                                      3660
ttegtetgat gttataatat taatattact gaateattgg ttttgeetge accteacaga
                                                                      3720
aatgttactg tgtcactttt ccctcgggag gaaattgttt ggtaatatag aaaggtgtat
                                                                      3780
gcaagttgaa tttgctgact ccagcacagt taaaaggtca atattctttt gacctgatta
                                                                      3840
atcagtcaga aagtccctat aggatagagc tggcagctga gaaattttaa aggtaattga
                                                                      3900
                                                                      3960
taattagtat ttgtaacttt ttaaagggct ctttgtatag cagaggatct catttgactt
tgttttgatg agggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt
                                                                      4020
gtccatgcag attttattgg cagctgtttt attgccattc aactagggaa atgaagaaat
                                                                      4080
cacgcagcct tttggttaaa tggcagtcaa aattttcctc agtgtattta gtgtgttcag
                                                                      4140
tgatgatatc actggttccc aactagatgc ttgttggcca cgggaaggga aatgacttgt
                                                                      4200
tctaattcta ggttcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga
                                                                      4260
cggtatttat gaatcagggt taggctccat atttaaagat agagccagtt tttttttaa
                                                                      4320
atagaaccca aattgtgtaa aaatgttaat tgggtttttt aaacattgtt ttatcaagtc
                                                                      4380
actgttaagt agaagaaagc catggtaaac tgatacataa cctaaattat aaaagcagaa
                                                                      4440
acctaactca ctcgtcaagg gaagttacct tttgaggaaa gttaaagtac ttttttccct
                                                                      4500
                                                                      4560
atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt
tatatcaaat cagaatgtaa acatgaactc ttgcatatat ttaaaattgt gttggaacat
                                                                      4620
                                                                      4680
ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnggatta tcattatgtg
anactggcag attgcagtgc ancettatge caataaaatg taatttaaca geeecagata
                                                                      4740
ttgttgaata ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt
                                                                      4800
                                                                      4860
ttttcttttt ttttcttttt cttttgtttc cttggtacta attttaattt ttatttggaa
gggagcagta taaagcttat ttgtatttag tagtgtatct catagataca gacaaggcaa
                                                                      4920
                                                                      4980
gagatgataa gctgtttaaa tagtgtttaa tattgattgg gggtggggag aaagaaaaag
tgtattactt aaagatacta tatacgtttt gtatatcatt aaatctttaa aagaaatnna
                                                                      5040
                                                                      5100
ataaatttat tgtttncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag
gatecetega ggggeee
                                                                      5117
```

```
<210> 24
<211> 1224
<212> PRT
<213> Homo Sapiens
<220>
<221> UNSURE
<222> (794)...(794)
<223> UNKNOWN
```

<400> 24 Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys 10 Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp 25 30 20 Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu 40 35 Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala 55 60 Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys 70 75 Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr 85

-34-

								•	-74-						
Cys	Thr	Gln	Met 100	Leu	Leu	Ser	Asp	Met 105	Trp	Met	Gly	Arg	Leu 110	Asn	Met
Arg	Lys	Asn 115	Ser	Trp	Tyr	Lys	Val 120	Ile	Leu	Ser	Ile	Leu 125	Val	Pro	Pro
Ala	Ile 130	Leu	Leu	Leu	Glu	Tyr 135	Lys	Thr	Lys	Ala	Glu 140	Met	Ser	His	Ile
Pro 145	Gln	Ser	Gln	Asp	Ala 150	His	Gln	Met	Thr	Met 155	Asp	Asp	Ser	Glu	Asn 160
Asn	Phe	Gln	Asn	Ile 165	Thr	Glu	Glu	Ile	Pro 170	Met	Glu	Val	Phe	Lys 175	Glu
Val	Arg	Ile	Leu 180	Asp	Ser	Asn	Glu	Gly 185	Lys	Asn	Glu	Met	Glu 190	Ile	Gln
	-	195	Lys	-			200		_	-		205			
	210		Ile			215					220				
225			Leu	-	230					235					240
			Gln	245					250					255	
		_	Val 260					265					270		
		275	Lys				280					285			
	290		Ile			295					300				
305			Phe		310					315					320
			Tyr	325					330	-				335	
-			Ala 340					345			_		350		
		355	Val				360					365			
	370		Ser		_	375		_	_		380				-
385			Ser	_	390			_	-	395					400
			Phe	405					410					415	
	-		Val 420					425	_				430		
		435	Gln				440					445			
	450		Ile			455					460				
465			Ile		470					475					480
			Lys	485					490					495	
			Leu 500					505					510		
		515	Gly				520					525			
	530		Phe			535					540				
545			Phe		550					555					560
GIU	Arg	val	Glu	Gln 565	Met	Cys	TTE	GIN	570	гàг	GLU	vaı	стА	575	Arg

-35-Val Asn Tyr Ile Lys Arg Ser Leu Gln Ser Leu Asp Ser Gln Ile Gly His Leu Gln Asp Leu Ser Ala Leu Thr Val Asp Thr Leu Lys Thr Leu Thr Ala Gln Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn Leu Ile Asp Asp Gly Pro Val Arg Pro Ser Val Trp Lys Lys His Gly Val Val Asn Thr
645 650 655 Leu Ser Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn Asn Pro Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln Cys Asn Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys Ile 725 730 735 Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro His 740 745 750 Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln Pro . 765 Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Xaa Glu Phe Gly Ala Phe Val 785 790 795 800 Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn Tyr 900 905 910 Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys Glu Glu Met Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr Trp 980 985 990 Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu Asp 1010 1015 Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Gln Arg Ala Ala 1030 1035 Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile Pro

-36-

Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala 1065 Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys 1075 1080 1085 Tyr Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu 1095 1100 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly 1105 1110 1115 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp 1125 1130 1135 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe 1140 1145 1150 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys 1160 1165 1155 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys 1170 1175 1180

1170 1175 1180

Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro 1185 1190 1195 120

Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser 1205 1210 1215

Ala Asn Ser Val Arg Leu Met Leu 1220

<210> 25

<211> 2180

<212> DNA

<213> Homo Sapiens

<400> 25

tegaggeeaa gaatteggea egagggeete gggeaggeee eetggagega eetgettett tgggcactgt tgctgaacag ggcacagatg gccatgtact tctgggagat gggttccaat 120 180 gcagtttcct cagctcttgg ggcctgtttg ctgctccggg tgatggcacg cctggagcct 240 gacgctgagg aggcagcacg gaggaaagac ctggcgttca agtttgaggg gatgggcgtt 300 gacctctttg gcgagtgcta tcgcagcagt gaggtgaggg ctgcccgcct cctcctccgt 360 cgctgcccgc tctgggggga tgccacttgc ctccagctgg ccatgcaagc tgacgcccgt 420 gccttctttg cccaggatgg ggtacagtct ctgctgacac agaagtggtg gggagatatg gccagcacta cacccatctg ggccctggtt ctcgccttct tttgccctcc actcatctac 480 540 accegectea teacetteag gaaateagaa gaggageeca caegggagga getagagttt 600 gacatggata gtgtcattaa tggggaaggg cctgtcggga cggcggaccc agccgagaag acgccgctgg gggtcccgcg ccagtcgggc cgtccgggtt gctgcggggg ccgctgcggg 660 720 gggcgccggt gcctacgccg ctggttccac ttctggggcg cgccggtgac catcttcatg 780 ggcaacgtgg tcagctacct gctgttcctg ctgcttttct cgcgggtgct gctcgtggat 840 ttccagccgg cgccgccgg ctccctggag ctgctgctct atttctgggc tttcacgctg ctgtgcgagg aactgcgcca gggcctgagc ggaggcgggg gcagcctcgc cagcgggggc 900 960 ecegggeetg gecatgeete actgageeag egeetgegee tetacetege egacagetgg aaccagtgcg acctagtggc tctcacctgc ttcctcctgg gcgtgggctg ccggctgacc 1020 1080 eegggtttgt accaectggg eegcactgte etetgeateg actteatggt titeaeggtg 1140 eggetgette acatetteae ggteaacaaa eagetgggge eeaagategt eategtgage aagatgatga aggacgtgtt cttcttcctc ttcttcctcg gcgtgtggct ggtagcctat 1200 ggcgtggcca cggagggct cctgaggcca cgggacagtg acttcccaag tatcctgcgc 1260 cgcgtcttct accgtcccta cctgcagatc ttcgggcaga ttccccagga ggacatggac 1320 gtggccctca tggagcacag caactgctcg tcggagcccg gcttctgggc acaccctcct 1380 ggggcccagg cgggcacctg cgtctcccag tatgccaact ggctggtggt gctgctcctc 1440 gtcatcttcc tgctcgtggc caacatcctg ctggtcaact tgctcattgc catgttcagt 1500 tacacattcg gcaaagtaca gggcaacagc gatetetact ggaaggegca gegttacege 1560 1620 ctcatccggg aattccactc tcggcccgcg ctggcccgc cctttatcgt catctcccac 1680 ttgcgcctcc tgctcaggca attgtgcagg cgacccsgga gcccccagcc gtcctccccg 1740 gccctcgagc atttccgggt ttacctttct aaggaagccg agcggaagct gctaacgtgg 1800 gaatcggtgc ataaggaqaa ctttctqctq qcacqcqcta gggacaagcg ggagagcgac 1860 tecgagmgte tgaagegeae gteceagaag gtggaettgg caetgaaaca getgggaeae

```
-37-
atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg
                                                                    1920
                                                                    1980
geocegeae etggtggeet tgteettgag gtgageecca tgteeatetg ggeeaetgte
aggaccacct ttgggagtgt catccttaca aaccacagca tgcccggctc ctcccagaac
                                                                    2040
cagteccage etgggaggat caaggeetgg atccerggee gttatecate tggaggetge
                                                                    2100
                                                                    2160
agggtccttg gggtaacagg gaccacagac ccctcaccac tcacagattc ctcacactgg
ggaaataaag ccatttcaga
                                                                    2180
     <210> 26
     <211> 725
      <212> PRT
     <213> Homo Sapiens
     <220>
     <221> UNSURE
     <222> (553)...(553)
     <223> UNKNOWN
     <221> UNSURE
     <222> (603)...(603)
     <223> UNKNOWN
     <400> 26
Ser Arg Pro Arg Ile Arg His Glu Gly Leu Gly Gln Ala Pro Trp Ser
                                   10
Asp Leu Leu Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met
           20
                               25
Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala
      35
                       40
                                              4.5
Cys Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu
                       55
Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
                   70
                                      75
Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg
               85
                                  90
Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln
          100
                              105
                                                  110
Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val
       115
                           120
                                              125
Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr
                      135
                                          140
Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr
                   150
                                      155
Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu
               165
                                   170
                                                     175
Glu Leu Glu Phe Asp Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val
           180
                               185
                                                  190
Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln
       195
                          200
                                              205
Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys
   210
                      215
                                          220
Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met
                  230
                                      235
Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Phe Ser Arg Val
               245
                                   250
                                                      255
Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu
           260
                              265
Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly
                           280
                                               285
Leu Ser Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly
```

His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp · 310 315 Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly 325 330 Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys 345 · 340 350 Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val 365 355 360 Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys . 375 380 Asp Val Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr 390 385 395 Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro 410 405 415 Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly 430 425 420 Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn 440 445 435 Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala 455 460 Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu 475 470 Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile 485 490 495 Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu 500 505 510 Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg 515 520 525 Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu 535 540 Leu Arg Gln Leu Cys Arg Arg Pro Xaa Ser Pro Gln Pro Ser Ser Pro 550 555 Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys 565 570 575 Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg 580 585 590 Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Xaa Leu Lys Arg Thr Ser 600 605 595 Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr 615 620 Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Thr Ser 630 635 625 Ala Pro Ala Pro Gly Gly Leu Val Leu Glu Val Ser Pro Met Ser Ile 645 650 Trp Ala Thr Val Arg Thr Thr Phe Gly Ser Val Ile Leu Thr Asn His 660 665 670 Ser Met Pro Gly Ser Ser Gln Asn Gln Ser Gln Pro Gly Arg Ile Lys 675 680 685 Ala Trp Ile Pro Gly Arg Tyr Pro Ser Gly Gly Cys Arg Val Leu Gly
690 700 700 695 Val Thr Gly Thr Thr Asp Pro Ser Pro Leu Thr Asp Ser Ser His Trp 705 710 715 Gly Asn Lys Ala Ile

<210> 27

<211> 7419

<212> DNA

<213> Homo Sapiens

<400> 27 60 cggggaccga tccagcctcc ggactctage ctaggctttt gcaaaaagct atttaggtga 120 cactatagaa ggtacgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc 180 qcaqccccqt cgccggcgga ggcggggcgcg ggcgcgtncc ctgtggccag tcacccggag gagttggtcg cacaattatg aaagactcgg cttctgctgc tagcgccgga gctgagttag 240 ttctgagaag gtttccctgg gcgttccttg tccggcggcc tctgctgccg cctccggaga 300 360 cqcttcccqa tagatggcta caggccgcgg aggaggagga ggtggagttg ctgcccttcc 420 ggagtccgcc ccgtgaggag aatgtcccag aaatcctgga tagaaagcac tttgaccaag agggaatgtg tatatattat accaagttcc aaggaccctc acagatgcct tccaggatgt 480 540 caaatttgtc agcaactcgt caggtgtttt tgtggtcgct tggtcaagca acatgcttgt tttactgcaa gtcttgccat gaaatactca gatgtgaaat tgggtgacca ttttaatcag 600 660 gtcataaatt ttcaaggggg ttctcattcc tacagagcta agtatgtgag gctatcatat 720 780 gacaccaaac ctgaagtcat tctgcaactt ctgcttaaag aatggcaaat ggagttaccc aaacttgtta tototgtaca tgggggcatg cagaaatttg agottcaccc acgaatcaag 840 900 cagttgcttg gaaaaggtct tattaaagct gcagttacaa ctggagcctg gattttaact 960 ggaggagtaa acacaggtgt ggcaaaacat gttggagatg ccctcaaaga acatgcttcc 1020 agatcatctc gaaagatttg cactatcgga atagctccat ggggagtgat tgaaaacaga aatgatettg ttgggagaga tgtggttget cettateaaa eettattgaa eeceetgage 1080 aaattgaatg ttttgaataa tctgcattcc catttcatat tggtggatga tggcactgtt 1140 1200 ggaaagtatg gggcggaagt cagactgaga agagaacttg aaaaaactat taatcagcaa agaattcatg ctaggattgg ccagggtgtc cctgtggtgg cacttatatt tgagggtggg 1260 ccaaatgtta tcctcacagt tcttgaatac cttcaggaaa gccccctgt tccagtagtt 1320 gtgtgtgaag gaacaggcag agctgcagat ctgctagcgt atattcataa acaaacagaa 1380 1440 gaaggaggga atcttcctga tgcagcagag cccgatatta tttccactat caaaaaaaca tttaactttg gccagaatga agcacttcat ttatttcaaa cactgatgga gtgcatgaaa 1500 agaaaggagc ttatcactgt tttccatatt gggtcagatg aacatcaaga tatagatgta 1560 gcaatactta ctgcactgct aaaaggtact aatgcatctg catttgacca gcttatcctt 1620 1680 acattggcat gggatagagt tgacattgcc aaaaatcatg tatttgttta tggacagcag 1740 tggctggttg gatccttgga acaagctatg cttgatgctc ttgtaatgga tagagttgca 1800 tttgtaaaac ttcttattga aaatggagta agcatgcata aattccttac cattccgaga 1860 ctggaagaac tttacaacac taaacaaggt ccaactaatc caatgctgtt tcatcttgtt cgagacgtca aacagggaaa tcttcctcca ggatataaga tcactctgat tgatatagga 1920 1980 cttgttattg aatatctcat gggaggaacc tacagatgca cctatactag gaaacgtttt cgattaatat ataatagtct tggtggaaat aatcggaggt ctggccgaaa tacctccagc 2040 2100 agcactcctc agttgcgaaa gagtcatgaa tcttttggca atagggcaga taaaaaggaa aaaatgaggc ataaccattt cattaagaca gcacagccct tccgaccaaa gattgataca 2160 2220 gttatggaag aaggaaagaa gaaaagaacc aaagatgaaa ttgtagacat tgatgatcca 2280 gaaaccaagc gctttcctta tccacttaat gaacttttaa tttgggcttg ccttatgaag 2340 aggcaggtca tggcccgttt tttatggcaa catggtgaag aatcaatggc taaagcatta 2400 gttgcctgta agatctatcg ttcaatggca tatgaagcaa agcagagtga cctggtagat 2460 gatacttcag aagaactaaa acagtattcc aatgattttg gtcagttggc cgttgaatta ttagaacagt ccttcagaca agatgaaacc atggctatga aattgctcac ttatgaactg 2520 2580 aagaactgga gtaattcaac ctgccttaag ttagcagttt cttcaagact tagacctttt 2640 gtagctcaca cctgtacaca aatgttgtta tctgatatgt ggatgggaag gctgaatatg aggaaaaatt cctggtacaa ggtcatacta agcattttag ttccacctgc catattgctg 2700 2760 ttagagtata aaactaaggc tgaaatgtcc catatcccac aatctcaaga tgctcatcag 2820 atgacaatgg atgacagcga aaacaacttt cagaacataa cagaagagat ccccatggaa 2880 gtgtttaaag aagtacggat tttggatagt aatgaaggaa agaatgagat ggagatacaa 2940 atgaaatcaa aaaagcttcc aattacgcga aagttttatg ccttttatca tgcaccaatt 3000 gtaaaattct ggtttaacac gttggcatat ttaggatttc tgatgcttta tacatttgtg 3060 gttcttgtac aaatggaaca gttaccttca gttcaagaat ggattgttat tgcttatatt 3120 tttacttatg ccattgagaa agtccgtgag atctttatgt ctgaagctgg gaaagtaaac cagaagatta aagtatggtt tagtgattac ttcaacatca gtgatacaat tgccataatt 3180 tettettea ttggatttgg actaagattt ggagcaaaat ggaactttge aaatgeatat 3240 gataatcatg tttttgtggc tggaagatta atttactgtc ttaacataat attttggtat 3300 3360 gtgcgtttgc tagattttct agctgtaaat caacaggcag gaccttatgt aatgatgatt ggaaaaatgg tggccaatat gttctacatt gtagtgatta tggctcttgt attacttagt 3420 tttggtgttc ccagaaaggc aatactttat cctcatgaag caccatcttg gactcttgct 3480 aaagatatag tttttcaccc atactggatg atttttggtg aagtttatgc atacgaaatt 3540

gatgtgtgtg	caaatgattc	tgttatccct	caaatctgtg	gtcctgggac	gtggttgact	3600
	aagcagtcta					3660
gcatttttca	acaatgtgta	tttacaaqtq	aaggcaattt	ccaatattgt	atggaagtac	3720
	attttattat					3780
	atatagtttc			_		3840
-	gaccaaaact		-			3900
	gtgttgaaat			_	_	3960
	ttcqtqtcac					4020
	gtgtcaacta	-			-	4080
	atctttcagc	-		_		4140
	ctagcaaagt		-		-	4200
	accttattga	_				4260
	cacttagctc					4320
	-					4380
-	ttttaatgaa		-	_		
	taccccagag	_				4440
-	caagtgctgt					4500
	tatttaataa			_		4560
_	caccaaccaa					4620
	ctggaaccaa	-		_		4680
-	ttggagcatt					4740
	acaagataaa					4800
	ctcttgctgg					4860
aaacaagcag	aaaaaatcag	tagaaggcca	tctaccgaag	acactcatga	agtagattcc	4920
aaagcagctt	taataccgga	ttggttacaa	gatagaccat	caaacagaga	aatgccatct	4980
gaagaaggaa	cattaaatgg	tctcacttct	ccatttaagc	cagctatgga	tacaaattac	5040
tattattcag	ctgtggaaag	aaataacttg	atgaggttat	cacagagcat	tccatttaca	5100
cctgtgcctc	caagagggga	gcctgtcaca	gtgtatcgtt	tggaagagag	ttcacccaac	5160
atactaaata	acagcatgtc	ttcttggtca	caactaggcc	tctgtgccaa	aatagagttt	5220
ttaagcaaag	aggagatggg	aggaggttta	cgaagagctg	tcaaagtaca	gtgtacctgg	5280
tcagaacatg	atatcctcaa	atcagggcat	ctttatatta	tcaaatcttt	tcttccagag	5340
gtggttaata	catggtcaag	tatttataaa	gaagatacag	ttctgcatct	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacgtttg	cctttaatca	aatgaaaccc	5460
	catattctcc					5520
	ttgctgtgga				_	5580
	agattattcc					5640
	aatatacaag					5700
	acccatctgt					5760
	atctaggaga	_		-		5820
	gaaagcttaa				_	5880
	ctcaggatga	_		_	-	5940
	caactaattc					6000
	gcacctcaca		_		_	6060
	agaaaggtgt					6120
	ttgacctgat			_		6180
	aaaggtaatt	_				6240
	ctcatttgac					6300
ataccattaa	ccaaaggtag	atatecatae	agattttatt	gacagetatt	ttattgccat	6360
	aaatgaagaa					6420
						6480
	tagtgtgttc gaaatgactt					6540
000	tttcaagagg	•				6600
					_	6660
	tttttttttttttttttttttttttt					6720
-	_	_			-	6780
	ataaaagcag					
	actttttcc					6840
	tttattgatt					6900
	gtgttggaac					6960
	tatcattatg					7020
	cagccccaga					7080
ctaagtttta	tgctttaatt	ttttttcttt	ttttttcttt	ttcttttgtt	tccttggtac	7140

PCT/US99/29996 WO 00/40614 -41-

aaaaagggcg gcccctctag aggatccctc gaggggccc

taattttaat ttttatttgg aagggagcag tataaagctt atttgtattt agtagtgtat ctcatagata cagacaaggc aagagatgat aagctgttta aatagtgktt aatattgatt qqqqqtqqqq agaaagaaaa agtgtattac ttaaagatac tatatacskt ttktatatca ttaaatcttt aaaagaaatn naataaattt attgtttnca aaaaaaaaac ccnntaaaaa

<210> 28 <211> 1865 <212> PRT

<213> Homo Sapiens

<400> 28 Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys 5. Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val 4.5 Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp Val Lys Leu Gly Asp His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn Phe Gln Gly Gly Ser His Ser Tyr Arg Ala Lys Tyr Val Arg Leu Ser Tyr Asp Thr Lys Pro Glu Val Ile Leu Gln Leu Leu Leu Lys Glu Trp Gln Met Glu Leu Pro Lys Leu Val Ile Ser Val His Gly Gly Met Gln Lys Phe Glu Leu His Pro Arg Ile Lys Gln Leu Leu Gly Lys Gly Leu Ile Lys Ala Ala Val Thr Thr Gly Ala Trp Ile Leu Thr Gly Gly Val Asn Thr Gly Val Ala Lys His Val Gly Asp Ala Leu Lys Glu His Ala Ser Arg Ser Ser Arg Lys Ile Cys Thr Ile Gly Ile Ala Pro Trp Gly Val Ile Glu Asn Arg Asn Asp Leu Val Gly Arg Asp Val Val Ala Pro Tyr Gln Thr Leu Leu Asn Pro Leu Ser Lys Leu Asn Val Leu Asn Asn Leu His Ser His Phe Ile Leu Val Asp Asp Gly Thr Val Gly Lys Tyr Gly Ala Glu Val Arg Leu Arg Arg Glu Leu Glu Lys Thr Ile Asn Gln Gln Arg Ile His Ala Arg Ile Gly Gln Gly Val Pro Val Val Ala Leu Ile Phe Glu Gly Gly Pro Asn Val Ile Leu Thr Val Leu Glu Tyr Leu Gln Glu Ser Pro Pro Val Pro Val Val Val Cys Glu Gly Thr Gly Arg Ala Ala Asp Leu Leu Ala Tyr Ile His Lys Gln Thr Glu Glu Gly Gly Asn Leu Pro Asp Ala Ala Glu Pro Asp Ile Ile Ser Thr Ile Lys Lys Thr Phe Asn Phe Gly Gln Asn Glu Ala Leu His Leu Phe Gln Thr Leu

-42-Met Glu Cys Met Lys Arg Lys Glu Leu Ile Thr Val Phe His Ile Gly · 375 Ser Asp Glu His Gln Asp Ile Asp Val Ala Ile Leu Thr Ala Leu Leu Lys Gly Thr Asn Ala Ser Ala Phe Asp Gln Leu Ile Leu Thr Leu Ala · 405 Trp Asp Arg Val Asp Ile Ala Lys Asn His Val Phe Val Tyr Gly Gln Gln Trp Leu Val Gly Ser Leu Glu Gln Ala Met Leu Asp Ala Leu Val Met Asp Arg Val Ala Phe Val Lys Leu Leu Ile Glu Asn Gly Val Ser Met His Lys Phe Leu Thr Ile Pro Arg Leu Glu Glu Leu Tyr Asn Thr Lys Gln Gly Pro Thr Asn Pro Met Leu Phe His Leu Val Arg Asp Val Lys Gln Gly Asn Leu Pro Pro Gly Tyr Lys Ile Thr Leu Ile Asp Ile 500 505 510 Gly Leu Val Ile Glu Tyr Leu Met Gly Gly Thr Tyr Arg Cys Thr Tyr Thr Arg Lys Arg Phe Arg Leu Ile Tyr Asn Ser Leu Gly Gly Asn Asn Arg Arg Ser Gly Arg Asn Thr Ser Ser Ser Thr Pro Gln Leu Arg Lys Ser His Glu Ser Phe Gly Asn Arg Ala Asp Lys Lys Glu Lys Met Arg His Asn His Phe Ile Lys Thr Ala Gln Pro Phe Arg Pro Lys Ile Asp Thr Val Met Glu Glu Gly Lys Lys Lys Arg Thr Lys Asp Glu Ile Val Asp Ile Asp Asp Pro Glu Thr Lys Arg Phe Pro Tyr Pro Leu Asn Glu Leu Leu Ile Trp Ala Cys Leu Met Lys Arg Gln Val Met Ala Arg Phe Leu Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr Cys Thr Gln Met Leu Leu Ser Asp Met Trp Met Gly Arg Leu Asn Met Arg Lys Asn Ser Trp Tyr Lys Val Ile Leu Ser Ile Leu Val Pro Pro Ala Ile Leu Leu Glu Tyr Lys Thr Lys Ala Glu Met Ser His 770 780 Ile Pro Gln Ser Gln Asp Ala His Gln Met Thr Met Asp Asp Ser Glu Asn Asn Phe Gln Asn Ile Thr Glu Glu Ile Pro Met Glu Val Phe Lys Glu Val Arg Ile Leu Asp Ser Asn Glu Gly Lys Asn Glu Met Glu Ile Gln Met Lys Ser Lys Lys Leu Pro Ile Thr Arg Lys Phe Tyr Ala Phe

-43-Tyr His Ala Pro Ile Val Lys Phe Trp Phe Asn Thr Leu Ala Tyr Leu · 855 Gly Phe Leu Met Leu Tyr Thr Phe Val Val Leu Val Gln Met Glu Gln 870 875 Leu Pro Ser Val Gln Glu Trp Ile Val Ile Ala Tyr Ile Phe Thr Tyr 885 890 Ala Ile Glu Lys Val Arg Glu Ile Phe Met Ser Glu Ala Gly Lys Val 900 905 910 Asn Gln Lys Ile Lys Val Trp Phe Ser Asp Tyr Phe Asn Ile Ser Asp 915 920 925 Thr Ile Ala Ile Ile Ser Phe Phe Ile Gly Phe Gly Leu Arg Phe Gly 935 940 Ala Lys Trp Asn Phe Ala Asn Ala Tyr Asp Asn His Val Phe Val Ala 950 955 Gly Arg Leu Ile Tyr Cys Leu Asn Ile Ile Phe Trp Tyr Val Arg Leu 965 970 Leu Asp Phe Leu Ala Val Asn Gln Gln Ala Gly Pro Tyr Val Met Met 980 985 990 Ile Gly Lys Met Val Ala Asn Met Phe Tyr Ile Val Val Ile Met Ala 995 1000 1005 Leu Val Leu Leu Ser Phe Gly Val Pro Arg Lys Ala Ile Leu Tyr Pro 1010 1015 1020 His Glu Ala Pro Ser Trp Thr Leu Ala Lys Asp Ile Val Phe His Pro 1030 1035 Tyr Trp Met Ile Phe Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys 1045 1050 1055 Ala Asn Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu 1060 1065 1070 Thr Pro Phe Leu Gln Ala Val Tyr Leu Phe Val Gln Tyr Ile Ile Met 1075 1080 1085 Val Asn Leu Leu Ile Ala Phe Phe Asn Asn Val Tyr Leu Gln Val Lys 1090 1095 1100 Ala Ile Ser Asn Ile Val Trp Lys Tyr Gln Arg Tyr His Phe Ile Met 1110 1115 Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser 1125 1130 1135 His Ile Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp 1140 1145 1150 Lys Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys 1155 1160 1165 Lys Leu His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu 1175 1180 Lys Asp Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr 1190 1195 Phe Glu Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp 1205 1210 1215 Arg Val Asn Tyr Ile Lys Arg Ser Leu Gln Ser Leu Asp Ser Gln Ile 1220 1225 1230 Gly His Leu Gln Asp Leu Ser Ala Leu Thr Val Asp Thr Leu Lys Thr 1235 1240 1245 Leu Thr Ala Gln Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile 1250 1255 1260 Thr Arg Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn Leu Ile Asp 1270 1275 Asp Gly Pro Val Arg Pro Ser Val Trp Lys Lys His Gly Val Val Asn 1285 1290 1295 Thr Leu Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn Asn Pro 1300 1305 1310 Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln Cys Asn 1320

Ile Phe G			1335	,				1340	ŀ			
Phe Pro G		1.	350				1355	1				1360
Ser Pro P	•	1365				1370)				1375	•
Ile Phe A	sn Lys 1380		ln Lys		Gly 1385		Ser	Ser	Thr	Ser 1390		Pro
His Leu S	er Ser 395	Pro P	ro Thr	Lys 1400		Phe	Val	Ser	Thr 1405		Ser	Gln
Pro Ser C 1410	ys Lys	Ser H	is Leu 1415		Thr	Gly	Thr	Lys 1420		Gln	Glu	Thr
Val Cys S 1425	er Lys		hr Glu 430	Gly	Asp	Asn	Thr 1435		Phe	Gly	Ala	Phe 1440
Val Gly H	is Arg	Asp S 1445	er Met	Asp	Leu	Gln 1450		Phe	Lys	Glu	Thr 1455	
Asn Lys I	le Lys 1460	Ile L	eu Ser	Asn	Asn 1465	Asn		Ser	Glu	Asn 1470	Thr	Leu
Lys Arg V 1			eu Ala	Gly 1480		Thr	Asp	Cys	His 1485		Thr	Ser
Ile Pro V 1490		Ser L	ys Gln 1495		Glu	Lys	Ile	Ser 1500		Arg	Pro	Ser
Thr Glu A	sp Thr		lu Val 510	Asp	Ser	Lys	Ala 1515		Leu	Ile	Pro	Asp 1520
Trp Leu G	ln Asp			Asn	Arg	Glu 1530		Pro	Ser	Glu	Glu 1535	
Thr Leu A	sn Gly 1540		hr Ser	Pro	Phe 1545		Pro	Ala	Met	Asp 1550		Asn
Tyr Tyr T	yr Ser 555	Ala V	al Glu	Arg 1560		Asn	Leu	Met	Arg 1565		Ser	Gln
Ser Ile P 1570	ro Phe	Thr P	ro Val 1575		Pro	Arg	Gly	Glu 1580		Val	Thr	Val
Tyr Arg L 1585		1	590				1595	•				1600
Ser Trp S		1605				1610)				1615	5 .
Glu Glu M	1620)			1625	5				1630)	
	635			1640)				1645	5		
Ser Phe L 1650			165	5				1660)			
Asp Thr V 1665		1	670				1675	,				1680
Ala Gln L		1685				1690)				1695	5
Pro Tyr S	1700)			1709	5				1710)	
Ala Gly G 1	ln Trp 715	Phe A	la Val	Glu 1720		Суѕ	Met	Thr	Gly 1725	Glu 5	Phe	Arg
Lys Tyr A 1730	sn Asn	Asn A	sn Gly. 173!		Glu	Ile	Ile	Pro 1740		Asn	Thr	Leu
Glu Glu I 1745	le Met		la Phe 750	Ser	His	Trp	Thr 1755		Glu	Tyr	Thr	Arg 1760
Gly Glu L	eu Leu	Val L 1765	eu Asp	Leu	Gln	Gly 1770		Gly	Glu	Asn	Leu 177!	
Asp Pro S	er Val 1780	Ile L	ys Ala	Glu	Glu 1789		Arg	Ser	Cys	Asp 1790		Val
Phe Gly P 1			eu Gly	Glu 1800		Ala	Ile	Lys	Asn 1805		Arg	Ala

-45-

Lys His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu
1810

Lys Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu
1825

Pro Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu
1845

Ser Thr Asn Ser Val Arg Leu Met Leu
1860

1865

<210> 29 <211> 4061 <212> DNA <213> Homo Sapiens

<400> 29

ggtctggaag cagagccgc ggagggagcg ccggggccct gggctgcagg aggttgcggc 60 ggccgcggca gcatggtggt gccggagaag gagcagagct ggatccccaa gatcttcaag 120 aagaagacct gcacgacgtt catagttgac tccacagatc cgggagggac cttgtgccag 180 240 tgtgggcqcc cccqqaccqc ccaccccqca gtggccatgg aggatgcctt cggggcagcc 300 gtggtgaccg tgtgggacag cgatgcacac accacggaga agcccaccga tgcctacgga qaqctqqact tcacqqqqqc cqgccqcaag cacaqcaatt tcctccqqct ctctqaccqa 360 acqqatccag ctgcagttta tagtctggtc acacgcacat ggggcttccg tgccccgaac 420 ctggtggtgt cagtgctggg gggatcgggg ggccccgtcc tccagacctg gctgcaggac 480 ctgctgcgtc gtgggctggt gcgggctgcc cagagcacag gagcctggat tgtcactggg 540 ggtctgcaca cgggcatcgg ccggcatgtt ggtgtggctg tacgggacca tcagatggcc 600 agcactgggg gcaccaaggt ggtggccatg ggtgtggccc cctggggtgt ggtccggaat 660 agagacacce tcatcaacce caagggeteg tteeetgega ggtaceggtg gegeggtgae 720 ccggaggacg gggtccagtt tcccctggac tacaactact cggccttctt cctggtggac 780 840 gacggcacac acggctgcct ggggggcgag aaccgcttcc gcttgcgcct ggagtcctac 900 atotcacage agaagacggg cgtgggaggg actggaattg acatecetgt cetgeteete 960 ctgattgatg gtgatgagaa gatgttgacg cgaatagaga acgccaccca ggctcagctc 1020 ccatgtctcc tcgtggctgg ctcaggggga gctgcggact gcctggcgga gaccctggaa 1080 gacactetgg ceceagggag tgggggagee aggeaaggeg aageeegaga tegaateagg 1140 cqtttctttc ccaaagggga ccttgaggtc ctgcaggccc aggtggagag gattatgacc 1200 cggaaggagc tcctgacagt ctattcttct gaggatgggt ctgaggaatt cgagaccata gttttgaagg cccttgtgaa ggcctgtggg agctcggagg cctcagccta cctggatgag 1260 ctgcgtttgg ctgtggcttg gaaccgcgtg gacattgccc agagtgaact ctttcggggg 1320 qacatecaat ggeggteett ceatetegaa getteetta tggacgeett getgaatgac 1380 1440 eggeetgagt tegtgegett geteatttee caeggeetea geetgggeea etteetgace ccgatgcgcc tggcccaact ctacagcgcg gcgccctcca actcgctcat ccgcaacctt ttggaccagg cgtcccacag cgcaggcacc aaagccccag ccctaaaagg gggagctgcg 1500 1560 gageteegge eccetgaegt ggggeatgtg etgaggatge tgetggggaa gatgtgegeg 1620 1680 ccgaggtacc cctccggggg cgcctgggac cctcacccag gccagggctt cggggagagc 1740 atgtatetge teteggacaa ggccaceteg cegetetege tggatgetgg cetegggcag 1800 gcccctgga gcgacctgct tctttgggca ctgttgctga acagggcaca gatggccatg 1860 tacttctggg agatgggttc caatgcagtt tcctcagctc ttgggggcctg tttgctgctc 1920 cgggtgatgg cacgcctgga gcctgacgct gaggaggcag cacggaggaa agacctggcg ttcaagtttg aggggatggg cgttgacctc tttggcgagt gctatcgcag cagtgaggtg 1980 agggetgece geeteeteet eegtegetge eegetetggg gggatgeeae ttgeeteeag 2040 2100 ctggccatgc aaqctgacgc ccgtgccttc tttgcccagg atggggtaca gtctctgctg 2160 acacagaagt ggtggggaga tatggccagc actacaccca tctgggccct ggttctcgcc 2220 ttottttgcc ctccactcat ctacacccgc ctcatcacct tcaggaaatc agaagaggag 2280 cccacacggg aggagctaga gtttgacatg gatagtgtca ttaatgggga agggcctgtc gggacggcgg acccagccga gaagacgccg ctgggggtcc cgcgccagtc gggccgtccg 2340 ggttgctgcg ggggccgctg cggggggcgc cggtgcctac gccgctggtt ccacttctgg 2400 2460 ggcgcgccgg tgaccatctt catgggcaac gtggtcagct acctgctgtt cctgctgctt ttetegeggg tgetgetegt ggattteeag eeggegeege eeggeteeet ggagetgetg 2520 ctctatttct gggctttcac gctgctgtgc gaggaactgc gccagggcct gagcggaggc 2580 gggggcagcc tcgccagcgg gggccccggg cctggccatg cctcactgag ccagcgcctg 2640 cgcctctacc tegeogacag etggaaccag tgcgacctag tggeteteac etgetteete 2700

-46-

ctgggcgtgg	gctgccggct	gaccccgggt	ttgtaccacc	tgggccgcac	tgtcctctgc	2760
atcgacttca	tggttttcac	ggtgcggctg	cttcacatct	tcacggtcaa	caaacagctg	2820
gggcccaaga	tcgtcatcgt	gagcaagatg	atgaaggacg	tgttcttctt	cctcttcttc	2880
ctcggcgtgt	ggctggtagc	ctatggcgtg	gccacggagg	ggctcctgag	gccacgggac	2940
	caagtatcct					3000
cagattcccc	aggaggacat	ggacgtggcc	ctcatggagc	acagcaactg	ctcgtcggag	3060
cccggcttct	gggcacaccc	tcctggggcc	caggcgggca	cctgcgtctc	ccagtatgcc	3120
aactggctgg	tggtgctgct	cctcgtcatc	ttcctgctcg	tggccaacat	cctgctggtc	3180
aacttgctca	ttgccatgtt	cagttacaca	ttcggcaaag	tacagggcaa	cagcgatctc	3240
tactggaagg	cgcagcgtta	ccgcctcatc	cgggaattcc	actctcggcc	cgcgctggcc	3300
ccgcccttta	tcgtcatctc	ccacttgcgc	ctcctgctca	ggcaattgtg	caggcgaccc	3360
cggagccccc	agccgtcctc	cccggccctc	gagcatttcc	gggtttacct	ttctaaggaa	3420
gccgagcgga	agctgctaac	gtgggaatcg	gtgcataagg	agaactttct	gctggcacgc	3480
gctagggaca	agcgggagag	cgactccgag	cgtctgaagc	gcacgtccca	gaaggtggac	3540
ttggcactga	aacagctggg	acacatccgc	gagtacgaac	agcgcctgaa	agtgctggag	3600
cgggaggtcc	agcagtgtag	ccgcgtcctg	gggtgggtgg	ccgaggccct	gagccgctct	3660
gccttgctgc	ccccaggtgg	gccgccaccc	cctgacctgc	ctgggtccaa	agactgagcc	3720
ctgctggcgg	acttcaagga	gaagccccca	caggggattt	tgctcctaga	gtaaggctca	3780
tctgggcctc	ggcccccgca	cctggtggcc	ttgtccttga	ggtgagcccc	atgtccatct	3840
gggccactgt	caggaccacc	tttgggagtg	tcatccttac	aaaccacagc	atgcccggct	3900
cctcccagaa	ccagtcccag	cctgggagga	tcaaggcctg	gatcccgggc	cgttatccat	3960
ctggaggctg	cagggtcctt	ggggtaacag	ggaccacaga	cccctcacca	ctcacagatt	4020
cctcacactg	gggaaataaa	gccatttcag	aggaaaaaaa	a		4061

<210> 30 <211> 1214 <212> PRT

<213> Homo Sapiens

<400> 30

Met Val Val Pro Glu Lys Glu Gln Ser Trp Ile Pro Lys Ile Phe Lys 5 10 Lys Lys Thr Cys Thr Thr Phe Ile Val Asp Ser Thr Asp Pro Gly Gly 20 25 30Thr Leu Cys Gln Cys Gly Arg Pro Arg Thr Ala His Pro Ala Val Ala 35 40 45 Met Glu Asp Ala Phe Gly Ala Ala Val Val Thr Val Trp Asp Ser Asp 55 60 Ala His Thr Thr Glu Lys Pro Thr Asp Ala Tyr Gly Glu Leu Asp Phe 65 70 75 80 Thr Gly Ala Gly Arg Lys His Ser Asn Phe Leu Arg Leu Ser Asp Arg 85 90 95 85 90 Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe 100 105 110 Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro 115 120 125 Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg 130 135 140 Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr 145 150 150 160 150 155 Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala 165 170 175 Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly 180 185 190 Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro 195 200 205 Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro 210 215 Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His 230

-47-Gly Cys Leu Gly Gly Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Trp Ala Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu

-48-

Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Glr 740 745 750 Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys 755 760 765 Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met	3
755 760 765 Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met	: •
770 775 780)
Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Phe Ser Arg Val 785 790 795 800	
Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu 805 810 815	
Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly 820 825 830	
Leu Ser Gly Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly 835 840 845	
His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp 850 855 860	
Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly 865 870 875 880	
Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys 885 890 895	i
Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val 900 905 910	
Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys 915 920 925	
Asp Val Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr 930 935 940	
Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro 945 950 955 960	}
Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly 965 970 975	
Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asm 980 985 990	
Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala 995 1000 1005	
Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu 1010 1015 1020	
Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile 1025 1030 1035 104	0
Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu 1045 1050 1055	1
Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg 1060 1065 1070	i
Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu 1075 1080 1085	l
Leu Arg Gln Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro 1090 1095 1100	•
Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys 1105 1110 1115 112	0:
Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg 1125 1130 1135	
Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser 1140 1145 1150	•
Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr 1155 1160 1165	:
Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg 1170 1175 1180	ſ
Val Leu Gly Trp Val Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro 1185 1190 1195 120	_

-49-

Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp 1205 1210

> <210> 31 <211> 4646 <212> DNA <213> Homo Sapiens

<400> 31

tegacecacg egteegeeca egegteegee caegegteeg eccaegegte egeecaegeg 60 teegeceaeg egteeggggt gaaagmramy emygektsms aaaaaeegte aettaggaaa 120 agatgteett tegggeagee aggeteagea tgaggaacag aaggaatgae actetggaea 180 240 gcacccggac cctgtactcc agcgcgtctc ggagcacaga cttgtcttac agtgaaagcg acttggtgaa ttttattcaa gcaaatttta agaaacgaga atgtgtcttc tttaccaaag 300 attccaaggc cacggagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag 360 420 gcacccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaattteeta 480 ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggaag tatatacgtc tgtcctgcga cacggacgcg gaaatccttt acgagctgct gacccagcac tggcacctga 540 aaacacccaa cetggtcatt tetgtgaccg ggggcgccaa gaacttegee etgaageege 600 660 gcatgcgcaa gatcttcagc cggctcatct acatcgcgca gtccaaaggt gcttggattc tcacgggagg cacccattat ggcctgatga agtacatcgg ggaggtggtg agagataaca 720 ccatcagcag gagttcagag gagaatattg tggccattgg catagcagct tggggcatgg 780 totocaaceg ggacaccoto atcaggaatt gcgatgotga gggctatttt ttagcccagt 840 900 accttatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt 960 tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaag ctccggaatc 1020 agctagagaa gtatatetet gagegeacta tteaagatte caactatggt ggeaagatee 1080 ccattgtgtg ttttgcccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca tcaaaaataa aattccttgt gtggtggtgg aaggctcggg ccagatcgct gatgtgatcg 1140 ctagectggt ggaggtggag gatgecetga catettetge egteaaggag aagetggtge 1200 gctttttacc ccgcacggtg tcccggctgc ctgaggagga gactgagagt tggatcaaat 1260 1320 ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaaaatg gaagaagctg gggatgaaat tgtgagcaat gccatctcct acgctctata caaagccttc agcaccagtg 1380 1440 agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg acttagccaa tgatgagatt ttcaccaatg accgccgatg ggagtctgct gaccttcaag 1500 1560 aagtcatgtt tacggctctc ataaaggaca gacccaagtt tgtccgcctc tttctggaga atggcttgaa cctacggaag tttctcaccc atgatgtcct cactgaactc ttctccaacc 1620 1680 acttcagcac gcttgtgtac cggaatctgc agatcgccaa gaattcctat aatgatgccc 1740 tcctcacgtt tgtctggaaa ctggttgcga acttccgaag aggcttccgg aaggaagaca 1800 gaaatggccg ggacgagatg gacatagaac tccacgacgt gtctcctatt actcggcacc 1860 ccctgcaagc tctcttcatc tgggccattc ttcagaataa gaaggaactc tccaaagtca tttgggagca gaccaggggc tgcactctgg cagccctggg agccagcaag cttctgaaga 1920 1980 ctctggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaatg agtacgagac ccgggctgtt gagctgttca ctgagtgtta cagcagcgat gaagacttgg 2040 2100 cagaacagct gctggtctat tcctgtgaag cttggggtgg aagcaactgt ctggagctgg cggtggaggc cacagaccag catttcatcg cccagcctgg ggtccagaat tttctttcta 2160 2220 agcaatggta tggagagatt tecegagaca ecaagaactg gaagattate etgtgtetgt ttattatacc cttggtgggc tgtggctttg tatcatttag gaagaaacct gtcgacaagc 2280 2340 acaagaaget getttggtae tatgtggegt tetteacete eccettegtg gtetteteet 2400 ggaatgtggt cttctacate geetteetee tgetgtttge etacgtgetg etcatggatt 2460 tocattoggt gccacacccc cccgagctgg toctgtactc gctggtcttt gtcctcttct 2520 gtgatgaagt gagacagtgg tacgtaaatg gggtgaatta ttttactgac ctgtggaatg tgatggacac gctggggctt ttttacttca tagcaggaat tgtatttcgg ctccactctt 2580 2640 ctaataaaag ctctttgtat tctggacgag tcattttctg tctggactac attatttca ctctaagatt gatccacatt tttactgtaa gcagaaactt aggacccaag attataatgc 2700 2760 tgcagaggat gctgatcgat gtgttcttct tcctgttcct ctttgcggtg tggatggtgg 2820 cctttggcgt ggccaggcaa gggatcctta ggcagaatga gcagcgctgg aggtggatat 2880 teegtteggt catetacgag eectacetgg ceatgttegg eeaggtgeee agtgaegtgg atggtaccac gtatgacttt gcccactgca ccttcactgg gaatgagtcc aagccactgt 2940 3000 gtgtggaget ggatgageae aacetgeece ggtteeeega gtggateaee ateeeeetgg 3060 tgtgcatcta catgttatcc accaacatcc tgctggtcaa cctgctggtc gccatgtttg

PCT/US99/29996 WO 00/40614

-50-

```
3120
qctacacqgt gggcaccgtc caggagaaca atgaccaggt ctggaagttc cagaggtact
tcctggtgca ggagtactgc agccgcctca atatcccctt ccccttcatc gtcttcgctt
                                                                      3180
acttctacat ggtggtgaag aagtgcttca agtgttgctg caaggagaaa aacatggagt
                                                                      3240
                                                                      3300
cttctgtctg ctgtttcaaa aatgaagaca atgagactct ggcatgggag ggtgtcatga
aggaaaacta ccttgtcaag atcaacacaa aagccaacga cacctcagag gaaatgaggc
                                                                      3360
                                                                      3420
atcgatttaġ acaactggat acaaagctta atgatctcaa gggtcttctg aaagagattg
ctaataaaat caaataaaac tgtatgaact ctaatggaga aaaatctaat tatagcaaga
                                                                      3480
                                                                      3540
tcatattaag gaatgctgat gaacaatttt gctatcgact actaaatgag agattttcag
acccctgggt acatggtgga tgattttaaa tcaccctagt gtgctgagac cttgagaata
                                                                      3600
                                                                      3660
aagtgtgtga ttggtttcat acttgaagac ggatataaag gaagaatatt tcctttatgt
                                                                      3720
gtttctccag aatggtgcct gtttctctct gtgtctcaat gcctgggact ggaggttgat
agtttaagtg tgttcttacc gcctcctttt tcctttaatc ttatttttga tgaacacata
                                                                      3780
tataggagaa catctateet atgaataaga acetggteat getttaetee tgtattgtta
                                                                      3840
                                                                      3900
ttttqttcat ttccaattga ttctctactt ttcccttttt tgtattatgt gactaattag
ttggcatatt gtwaaaagtc tctcaaatta ggccagattc taaaacatgc tgcagcaaga
                                                                      3960
ggaccccgct ctcttcagga aaagtgtttt catttctcag gatgcttctt acctgtcaga
                                                                      4020
ggaggtgaca aggcagtete ttgetetett ggaeteacea ggeteetatt gaaggaacea
                                                                      4080
                                                                      4140
cccccattcc taaatatgtg aaaagtcgcc caaaatgcaa ccttgaaagg cactactgac.
tttgttctta ttggatactc ctcttattta ttattttcc attaaaaata atagctggct
                                                                      4200
attatagaaa atttagacca tacagagatg tagaaagaac ataaattgtc cccattacct
                                                                      4260
taaggtaatc actgctaaca atttctggat ggtttttcaa gtctatttt tttctatgta
                                                                      4320
                                                                      4380
tgtctcaatt ctctttcaaa attttacaga atgttatcat actacatata tacttttat
gtaagctttt tcacttagta ttttatcaaa tatgttttta ttatattcat agccttctta
                                                                      4440
                                                                      4500
aacattatat caataattgc ataataggca acctctagcg attaccataa ttttgctcat
tgaaggctat ctccagttga tcattgggat gagcatcttt gtgcatgaat cctattgctg
                                                                      4560
                                                                      4620
tatttgggaa aattttccaa ggttagattc caataaatat ctatttatta ttaaaaaaaa
                                                                      4646
aaaaaaagg gcggccgctc tagagt
```

<210> 32 <211> 1104 <212> PRT

<213> Homo Sapiens

<400> 32

Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg Arg Asn Asp 5 10 Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser Arg Ser Thr 25 Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile Gln Ala Asn 40 Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser Lys Ala Thr 55 60 Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly 75 Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys 90 Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly 105 110 100 Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile 120 125 Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu 130 135 Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro Arg 150 155 Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser Lys Gly 170 165 Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met Lys Tyr Ile 185 190 180 Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser Glu Glu Asn 200

-51-Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Cys Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val Glu Gly Ser Gly Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg

Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu Phe Ile Ile Pro Leu 695 700 Val Gly Cys Gly Phe Val Ser Phe Arg Lys Lys Pro Val Asp Lys His 715 710 Lys Lys Leu Leu Trp Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val 730 · 725 Val Phe Ser Trp Asn Val Val Phe Tyr Ile Ala Phe Leu Leu Phe 745 740 Ala Tyr Val Leu Leu Met Asp Phe His Ser Val Pro His Pro Pro Glu 760 755 Leu Val Leu Tyr Ser Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg 780 775 Gln Trp Tyr Val Asn Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val 795 790 Met Asp Thr Leu Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg 810 815 805 Leu His Ser Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe 820 825 Cys Leu Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr 845 840 835 Val Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu 860 855 850 Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val Ala 870 875 Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln Arg Trp 890 885 Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu Ala Met Phe 905 900 Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr Asp Phe Ala His 920 915 Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu Cys Val Glu Leu Asp 935 940 Glu His Asn Leu Pro Arg Phe Pro Glu Trp Ile Thr Ile Pro Leu Val 955 950 Cys Ile Tyr Met Leu Ser Thr Asn Ile Leu Leu Val Asn Leu Leu Val 970 965 Ala Met Phe Gly Tyr Thr Val Gly Thr Val Gln Glu Asn Asn Asp Gln 980 985 Val Trp Lys Phe Gln Arg Tyr Phe Leu Val Gln Glu Tyr Cys Ser Arg 995 1000 1005 Leu Asn Ile Pro Phe Pro Phe Ile Val Phe Ala Tyr Phe Tyr Met Val 1010 1015 1020 Val Lys Lys Cys Phe Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser 1025 1030 1035 104 Ser Val Cys Cys Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu 1045 1050 Gly Val Met Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn 1060 1065 1070 Asp Thr Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys 1075 1080 1085 Leu Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys 1095

CORRECTED VERSION

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 13 July 2000 (13.07.2000)

PCT

(10) International Publication Number WO 00/40614 A3

- (51) International Patent Classification7: C07K 14/705, C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N 33/53, A61K 38/17
- (21) International Application Number: PCT/US99/29996
- (22) International Filing Date: 20 December 1999 (20.12.1999)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/114,220 30 December 1998 (30.12.1998) US 60/120,018 29 January 1999 (29.01.1999) US 60/140,415 22 June 1999 (22.06.1999) US

- (71) Applicant (for all designated States except US): BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 1 Deaconess Road, Boston, MA 02215 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): SCHARENBERG, Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA 02420 (US).

(74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

- (81) Designated States (national): AU, CA, JP, US.
- (84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

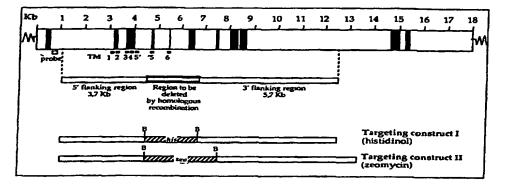
Published:

with international search report

- (88) Date of publication of the international search report: 22 February 2001
- (48) Date of publication of this corrected version: 30 August 2001
 - ction:
- (15) Information about Correction: see PCT Gazette No. 35/2001 of 30 August 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CHARACTERIZATION OF THE SOC/CRAC CALCIUM CHANNEL PROTEIN FAMILY



(57) Abstract: Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.



CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

Field of the Invention

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

Background of the Invention

5

10

15

20

25

30

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca²⁺ ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca²⁺ ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca²⁺ cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed store operated calcium influx, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

-2-Summary of the Invention

5

10

15

20

25

30

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca2+ flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca²⁺ from intracellular calcium stores, allowing Ca²⁺ influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25

30

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

- a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;

5

10

15

20

25

30

- c) isolating the SOC/CRAC molecule from the complex; and
- d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

5

10

15

20

25

30

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)2, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules in vivo or in vitro.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;
- b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

WO 00/40614 PCT/US99/29996
-6-

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

5

10

15

20

25

30

- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEO ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

- a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and
- b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

5

10

15

20

25

30

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, inter alia, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEQ ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

5

10

15

20

25

30

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a C. Elegans polypeptide at the c05c12.3 locus.

SEQ ID NO:14 is the amino acid sequence of a C. Elegans polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a C. Elegans polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

5

10

15

20

25

30

PCT/US99/29996 WO 00/40614 -9-

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEQ ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEQ ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEQ ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEO ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEO ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

Figure 1 is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca²⁺ flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca2+ from intracellular calcium stores, allowing Ca2+ influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

-10-

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

5

10

15

20

25

30

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca²⁺ fluxing and, in particular, lymphocyte Ca²⁺ fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

WO 00/40614 PCT/US99/29996 -11-

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

5

10

15

20

25

30

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid" refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca²⁺ transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

Science, 1992; 256:498-502; and Hoth. M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

- (a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25

30

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

5

10

15

20

25

30

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified in vitro by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

-14-

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

5

10

15

20

25

30

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEQ ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEO ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

-15-

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

5

10

15

20

25

30

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (ftp:/ncbi.nlm.nih.gov/pub/). Exemplary tools include the BLAST system available at http://www.ncbi.nlm.nih.gov. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVetor sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphoimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC cDNA clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

WO 00/40614 PCT/US99/29996 -16-

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

5

10

15

20

25

30

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

5

10

15

20

25

30

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

-18-

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

5

10

15

20

25

30

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., Nat. Med. 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In

addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5,SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or

5

10

15

20

25

30

experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred internucleoside phosphorothioates, alkylphosphonates, synthetic linkages are alkylphosphonothioates, phosphoramidates, phosphorodithioates, phosphate esters, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

5

10

15

20

25

30

-20-

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-Oalkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as <u>E.coli</u> and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

5

10

15

20

25

30

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed -22-

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

5

10

15

20

25

30

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the group consisting of SEQ ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

-23-

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

5

10

15

20

25

30

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1α, which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., <u>E. coli</u>), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

WO 00/40614 PCT/US99/29996
-24-

"knock-outs" in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

5

10

15

20

25

30

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca²⁺ fluxing, high selectivity, a unitary

WO 00/40614 PCT/US99/29996 -25-

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

5

10

15

20

25

30

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in Science 278:82-87, 1997, whereby proteins can be designed de novo. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

5

10

15

20

25

30

Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., <u>E. coli</u>, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

-27**-**

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

5

10

15

20

25

30

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of SOC/CRAC polypeptides, including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

WO 00/40614 PCT/US99/29996 -28-

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

5

10

15

20

25

30

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

WO 00/40614 PCT/US99/29996 -29-

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

10

15

20

25

30

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

5

10

15

20

25

30

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

WO 00/40614 PCT/US99/29996 -31-

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

5

10

15

20

25

30

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

-32-

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

5

10

15

20

25

30

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention. a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

5

10

15

20

25

30

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. Trans-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such trans-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic nonhuman animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

5

10

15

20

25

30

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca²⁺ fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

5

10

15

20

25

30

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca²⁺. Such separation is preferred so that a change in Ca²⁺ concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cellmembrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel Detection of a decrease in the activation and a given reference calcium concentration.

WO 00/40614 PCT/US99/29996 -36-

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5

10

15

20 -

25

30

SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or

polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include agents such as dihydropyridines (DHPs), phenylalkylamines, omega conotoxin (omega.-CgTx) and pyrazonoylguanidines.

10

15

20

25

30

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

5

10

15

20

25

30

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromotograpic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β -galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

5

10

15

20

25

30

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and Dictyostelium myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., Proc Natl Acad Sci U S A, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or Dictyostelium MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

-40-

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10⁷ M⁻¹, more preferably at least about 10⁸ M⁻¹, and most preferably at least about 10⁹ M⁻¹. The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

5

10

15

20

25

30

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced in vitro or in vivo in a host. Such techniques include transfection of nucleic acid-CaPO₄ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids

5

10

15

20

25

30

implantation.

Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sylastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, are used herein, means that the implant is constructed and arranged to delivery therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing ex vivo gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction in vitro of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. In vivo gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

10

15

20

25

30

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode C. elegans is nearing completion, and IP3-dependent calcium signals have recently been shown to be required for one or more aspects of C. elegans development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily excluding all proteins below approximately 1000 amino acids in size, followed by focusing on remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated CO5C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was noteable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

We then used C05C12.3 for BLAST alignment screening of the rest of the C. elegans genome and also mammalian databases for homologous proteins, revealing two other C. elegans homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and -43-

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRAC-Subsequent BLAST searches picked up mouse EST like currents were best defined. sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures

5

10

15

20

25

30

Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

-44-

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promers were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

5

10

15

20

25

30

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y., N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in <u>Capacitative Calcium Entry</u>, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2+ indicators with greatly improved fluorescence properties, <u>J. Biol</u>

-45-

5

10

15

20

25

30

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2⁺ indicators with greatly improved fluorescence properties, <u>J Biol Chem 260</u>, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

-46-

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., *EMBO J*, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (*Cell*, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure 1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

5

10

15

20

25

30

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

WO 00/40614 PCT/US99/29996
-47-

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in Xenopus oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: Pharmacologic Behavior of SOC/CRAC

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: Gene targeting

5

10

15

20

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectivelly refered to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

-48-Claims

1. An isolated nucleic acid molecule, comprising:

5

10

15

20

25

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEO ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).
- 2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.
- 3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.
- 4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.
- 5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.
- 6. An isolated nucleic acid molecule selected from the group consisting of
- (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,
 - (b) complements of (a),
- provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of
 - (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,
 - (2) complements of (1), and
 - (3) fragments of (1) and (2).

WO 00/40614 PCT/US99/29996

- 7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:
 - (1) at least two contiguous nucleotides nonidentical to the sequence group,
 - (2) at least three contiguous nucleotides nonidentical to the sequence group,
 - (3) at least four contiguous nucleotides nonidentical to the sequence group,
 - (4) at least five contiguous nucleotides nonidentical to the sequence group,
 - (5) at least six contiguous nucleotides nonidentical to the sequence group,
 - (6) at least seven contiguous nucleotides nonidentical to the sequence group.
- 8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.
- 9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.
 - 10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.
 - 11. A host cell transformed or transfected with the expression vector of claim 10.
- 20 12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.
 - 13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.
- 25 14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28. SEQ ID NO:30, and SEQ ID NO:32.

10

15

20

25

-50-15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3.

PCT/US99/29996

- 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.
- 16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.
- 17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEO ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.
 - 18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.
 - 19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.
 - 20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:
 - a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;
 - c) isolating the SOC/CRAC molecule from the complex; and
 - d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.
 - 21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.
 - 22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

- WO 00/40614 PCT/US99/29996 -51-
- 23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3. SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.
- 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

10

15

20

25

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;
- b) detecting a Ca²⁺ concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.
- 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:
- a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and
- b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.
- 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.
- 27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.
- 28. The method of claim 25, wherein the test sample is tissue.
- 29. The method of claim 25, wherein the test sample is a biological fluid.

WO 00/40614 PCT/US99/29996 -52-

- 30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).
- 31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to SOC/CRAC and polyclonal antisera to SOC/CRAC.
- 32. A kit, comprising a package containing:

5

10

15

20

25

30

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

- a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.
- 33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.
- 34. A pharmaceutical composition comprising:
- a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and
 - a pharmaceutically acceptable carrier.
- 35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.
- 36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:
- a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;
- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC kinase activity.
- 37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

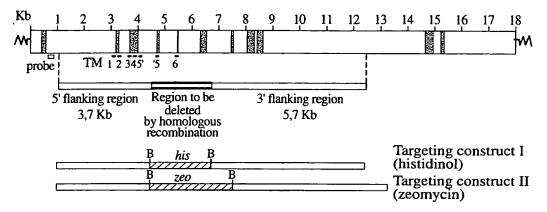


Fig. 1

-1-

SEQUENCE LISTING

```
<110> Beth Israel Deaconess Medical Center, Inc.
             Scharenberg, Andrew
      <120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY
      <130> B0662/7026WO/ERP/KA
      <150> U.S. 60/114,220
      <151> 1998-12-30
      <150> U.S. 60/120,018
      <151> 1999-01-29
      <150> U.S. 60/140,415
      <151> 1999-06-22
      <160> 32
      <170> FastSEO for Windows Version 3.0
      <210> 1
      <211> 1212
      <212> DNA
      <213> Homo Sapiens
      <400> 1
gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaag ccacttggaa
                                                                          60
actqqaacca aaqatcaaqa aactqtttgc tctaaagcta cagaaggaga taatacagaa
                                                                         120
tttggagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca
                                                                         180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt
                                                                         240
tctcttgctg gatttactga ctgtcacaga acttccattc ctgttcattc aaaacaagaa
                                                                         300
aaaatcaqta qaaqqccatc taccqaaqac actcatgaag tagattccaa agcagcttta
                                                                         360
                                                                         420
ataccggttt gtagatttca actaaacaga tatatattat taaatacatt aaactttttt
agataagatc tacaaagtgg tgatatttgg gactatatca aaaattcaaa aaaatttttc ttaagaaaac tgactttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc
                                                                         480
                                                                         540
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg
                                                                         600
aagctaatta tatgaactca tttccagcta actgcgataa tgattgattt tgcaaattcc
                                                                         660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt
                                                                         720
                                                                         780
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt
acattcatat atatctgttc tggcatagag agattatgtt gttaaaatca tgttattgaa
                                                                         840
                                                                         900
aaaagttatt tcagtgggga aagaggttag ttaacaaaga gattcacagt aacaaatcct
cctttctgga gggactcttc ctgaccctga gctgcacaac tttgcaacaa attaaagcct
                                                                         960
aaccgaagat gacctcacaa tggcaattta gaactcatgg gagtcaactt acataaacgg
                                                                        1020
                                                                        1080
tatttgattt ctgataagat agtggaatta ttggttatag atgacaaaat aagtatgttt
                                                                        1140
aaagtgatga tggacataaa aaagttttaa atataaaaca tgagaaaaga aggagatact
attcaaaaag actggcaaat ttgaaaaact agaaataaaa aaaaaaaaa aaaatgagcg
                                                                        1200
gccgcaagct tt
                                                                        1212
      <210> 2
      <211> 141
      <212> PRT
      <213> Homo Sapiens
      <400> 2
Ala Arg Gly Lys Phe Phe Val Ser Thr Pro Ser Gln Pro Ser Cys Lys
                                 10
                5
Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys
                                 25
            20
```

-2-Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe Val Gly His Arg 40 45 Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Lys 55 60 Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu Lys Arg Val Ser 75 Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser Ile Pro Val His 90 85 Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro Ser Thr Glu Asp Thr His 100 105 110 Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Val Cys Arg Phe Gln Leu 120 125 115 Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn Phe Phe Arg 135 <210> 3 <211> 739 <212> DNA <213> Homo Sapiens <220> <221> unsure <222> (5)...(5) <223> UNKNOWN <221> unsure <222> (21)...(22) <223> UNKNOWN <221> unsure <222> (29)...(29) <223> UNKNOWN <400> 3 60 tcgantaggg gtcttccacc nncatactng gatgatggtt ggtgaagtct atgcatacga 120 aattgatgtg tgtgcaaacg attctgttat ccctcaaatc tgtggtcctg ggacgtggtt gactccattt cttcaagcag tctacctctt tgwacagtat atcattatgg ttaatcttct 180 240 tattgcattt ytcaacaatg tgtatttaca agtgaaggca atttccaata ttgyatggaa gtaccagcgt tatcatttta ttatggctta tcatgagaaa ccagttctgc ctcctccact 300 360 tatcattctt agccatatag tttctctgtt ttgctgcata tgtaagagaa gaaagaaaga taagacttcc gatggaccaa aacttttctt aacagaagaa gatcaaaaga aacttcatga 420 480 ttttgaagag cagtgtttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg gagtgaagag agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa 540 600 agaagttgga gatccgtgtc aactacataa aaagatcatt acaatcatta gattctcaaa ttggccattt gcaagatctt tcagccctga cggtagatac attaaaaaca ctcactggcc 660 720 aaaaqcqtcq qaaqctaqca aagttcataa tgaaatcaca cgagaactga gcatttccaa 739 acacttggct caaaacctt <210> 4 <211> 235 <212> PRT <213> Homo Sapiens <220> <221> UNSURE <222> (41)...(41) <223> UNKNOWN <221> UNSURE

<222> (54)...(54)

<221> unsure <222> (375)...(375)

<223> g or c

WO 00/40614 PCT/US99/29996

-3-<223> UNKNOWN <221> UNSURE <222> (68)...(68) <223> UNKNOWN <400> 4 Met Met Val Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys Ala Asn 10 5 Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu Thr Pro 30 25 20 Phe Leu Gln Ala Val Tyr Leu Phe Xaa Gln Tyr Ile Ile Met Val Asn 45 35 40 Leu Leu Ile Ala Phe Xaa Asn Asn Val Tyr Leu Gln Val Lys Ala Ile 55 60 Ser Asn Ile Xaa Trp Lys Tyr Gln Arg Tyr His Phe Ile Met Ala Tyr 65 70 75 80 His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile 90 Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp Lys Thr 100 105 110100 105 Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu 125 . 120 His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu Lys Asp 140 135 Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr Phe Glu 155 150 Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp Pro Cys 165 170 175 Gln Leu His Lys Lys Ile Ile Thr Ile Ile Arg Phe Ser Asn Trp Pro 190 185 Phe Ala Arg Ser Phe Ser Pro Asp Gly Arg Tyr Ile Lys Asn Thr His 195 200 Trp Pro Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg 210 215 220 210 215 Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn 230 <210> 5 <211> 1579 <212> DNA <213> Homo Sapiens <220> <221> unsure <222> (368)...(368) <223> g or c <221> unsure <222> (372)...(372) <223> g or c <221> unsure <222> (374)...(374) <223> g or a

<221> unsure

```
<222> (387)...(387)
      <221> unsure
      <222> (482)...(482)
                                                                        60
acgtcgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc ggcatggtgt
tgtaaataca cttagctcct ctcttcctca aggtgatctt gaaagtaata atccttttca
                                                                       120
ttgtaatatt ttaatgaaag atgacaaaga tccccagtgt aatatatttg gtcaagactt
                                                                       180
acctgcagta ccccagagaa aagaatttaa ttttccagag gctggttcct cttctggtgc
                                                                       240
cttattccca agtgctgttt cccctccaga actgcgacag agactacatg gggtagaact
                                                                       300
cttaaaaata tttaataaaa atcaaaaatt aggcagttca tctactagca taccacatct
                                                                       360
qtcatccsca csarscaaat tttttgntag tacaccatct cagccaagtt gcaaaagcca
                                                                       420
cttggaaact ggaaccaaag atcaagaaac tgtttgctct aaagctacag aaggagataa
                                                                       480
                                                                       540
tncagaattt ggagcatttg taggacacag agatagcatg gatttacaga ggtttaaaga
aacatcaaac aagataaaaa tactatccaa taacaatact totgaaaaca otttgaaacg
                                                                       600
agtgagttct cttgctggat ttactgactg tcacagaact tccattcctg ttcattcaaa
                                                                       660
acaagaaaaa atcagtagaa ggccatctac cgaagacact catgaagtag attccaaagc
                                                                       720
                                                                       780
agctttaata ccggtttgta gatttcaact aaacagatat atattattaa atacattaaa
cttttttaga taagatctac aaagtggtga tatttgggac tatatcaaaa attcaaaaaa
                                                                       840
atttttctta agaaaactga ctttagcata gtagcagtta cagaaaagtt tcttacagtg
                                                                       900
aatagtcagg aattttaaag aaaaatttat gcagaataaa ggcaggaatc tctttttgtt
                                                                       960
tgaattgaag ctaattatat gaactcattt ccagctaact gcgataatga ttgattttgc
                                                                      1020
aaattccctt taaaagcaca cactgacaag acaaaaagct caggaaaagg cagaaaaatt
                                                                      1080
actcctttat aatcaagtat tatatataag tcagtgctca taattttgct caagaaaata
                                                                      1140
ttgacttaca ttcatatata tctgttctgg catagagaga ttatgttgtt aaaatcatgt
                                                                      1200
                                                                      1260
tattqaaaaa agttatttca gtggggaaag aggttagtta acaaagagat tcacagtaac
aaatcctcct ttctggaggg actcttcctg accctgagct gcacaacttt gcaacaaatt
                                                                      1320
aaagcctaac cgaagatgac ctcacaatgg caatttagaa ctcatgggag tcaacttaca
                                                                      1380
taaacggtat ttgatttctg ataagatagt ggaattattg gttatagatg acaaaataag
                                                                      1440
tatgtttaaa gtgatgatgg acataaaaaa gttttaaata taaaacatga gaaaagaagg
                                                                      1500
                                                                      1560
agatactatt caaaaagact ggcaaatttg aaaaactaga aataaaaaaa aaaaaaaaa
                                                                      1579
atgagcggcc gcaagcttt
      <210> 6
      <211> 243
      <212> PRT
      <213> Homo Sapiens
      <221> UNSURE
      <222> (103)...(105)
<223> UNKNOWN
      <221> UNSURE
      <222> (109)...(109)
      <223> UNKNOWN
      <221> UNSURE
      <222> (141)...(141)
      <223> UNKNOWN
      <400> 6
Val Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn
                                    10
Asn Pro Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln
                         25
                                                    30
Cys Asn Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu
                            40
                                                4.5
```

-5-Phe Asn Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser 55 Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu 75 70 Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser 90 85 95 Ile Pro His Leu Ser Ser Xaa Xaa Xaa Lys Phe Phe Xaa Ser Thr Pro 100 105 110 Ser Gln Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln 115 120 125 Glu Thr Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Xaa Glu Phe Gly 135 140 Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu 145 150 155 Thr Ser Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn 165 170 Thr Leu Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg 180 185 190 Thr Ser Ile Pro Val His Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro 200 205 Ser Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro . 210 215 220 Val Cys Arg Phe Gln Leu Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn 230 235 Phe Phe Arg <210> 7 <211> 3532

<212> DNA <213> Mus Musculus <220> <221> unsure <222> (2420)...(2420) <223> unknown <221> unsure <222> (2434)...(2434)

<223> unknown

<221> unsure

<222> (2461)...(2461)

<223> unknown

<221> unsure

<222> (2466)...(2466)

<223> unknown

<221> unsure

<222> (2470)...(2470)

<223> unknown

<400> 7

60 attatggett atcatgaaaa accagteetg ceteeteete ttateateet eageeatata gtttcactgt tttgctgtgt atgcaaaaga agaaagaaag ataagacttc cgatgggcca 120 aaacttttct taacagaaga agatcaaaag aaactccatg attttgaaga gcagtgtgtt 180 gagatgtact ttgatgagaa agatgacaaa ttcaattctg ggagtgaaga gagaatccgg 240 gtcacttttg aaagagtgga gcagatgagc attcagatta aagaagttgg agatcgtgtc 300 aactacataa aaagatcatt acagtcttta gattctcaaa ttggtcatct gcaagatctc 360

WO 00/40614	1				PCT/US	599/29996
			-6-			
			=	20222000		420
tcagccctaa	cagtagatac	accgaaaaca	actatate	agaaageeee	torgraph	480
aaagtgcaca	atgagatcac	acgagaaccg	agracticca	atactatasa	coagaacccc	540
actgatgatg	ttcctgtaag	teeeestage	gaagaaccta	ttatttata	tatttttata	600
tcctctctc	ctcaaggtga	tegggaaagt	trtagagaga	atttagggat	cattletate	660
	aagaccccca					720
	tcaacattcc					780
	cagaattacg					
	aattaggcag					840
	tgagtacccc					900
aaagatcaag	aacccatttt	ctataaagct	gcagaagggg	ataacataga	atttggagca	960
tttgtgggac	acagagatag	tatggactta	cagaggttta	aagaaacatc	aaacaaaata	1020
	tatctaatga					1080
	gttgtaagac					1140 1200
	cgacggaaga					
tggttacgag	atagaccatc	aaacagagaa	atgecatetg	aaggaggaac	actaaacggt	1260
	catttaagcc					1320
	tgaggttgtc					1380 1440
	tgtaccgtct					1500
tcatggtctc	agctaggcct	ctgtgccaaa	actgagtttt	taagtaaaga	ggaaatggaa	1560
ggtggtttac	gaagagcagt	caaagtgctg	tgtacctggt	cagagcacga	tateetgaag	
	tctatatcat					1620 1680
	aagatacggt					
	tcacatttgc					1740
aggttccttg	aagttttcct	gttgtactgc	catteageag	ggcagtggtt	tgetgtagaa	1800 1860
gagtgcatga	ctggtgaatt	tagaaaatac	aacaacaaca	acggcgacga	aatcattcct	1920
acaaatactc	tagaagagat	catgctagcc	tttagccact	ggacctatga	acacaccaga	1980
ggggagttac	tggtacttga	cttacaagga	grgggagaaa	actigactga	totaccigia	2040
ataaaagctg	aagaaaaaag	accetgigae	teenetates	ttattaatat	cctaggagaa	2100
gatgcaataa	aaaacttcaa	gagecaaaca	ceaetytaa	taaaattata	tttaataaa	2160
aacttcccag	atttgaagag	gaatgactac	acgeeeeega	caaaactata	gaagaaga	2220
	agatttgaat					2280
	tctgatgtta					2340
tgtaactgtc	agttttcctt	tegggggaat	cotattata	ttannana	cotcaaacya	2400
gettgetgge	cccacacata	gcccagaggc	atarcaasas	atazataata	teggaggig	2460
gaggetgeag	atgccagtgn	adagigetag	testatata	greagegere	20030005	2520
naaggneggn	accettgetg agtetteaga	tttaattaa	tastcaatca	cacceggege	aggaccacca	2580
accaaagtca	gcagcttttt	nantagatan	agtettagte	agtttatact	ctatgataat	2640
actabactyc	ctcaactagg	tatttataa	caccacacaca	ctactcctta	caattttact	2700
	gttacaaagc					2760
	agtattgagg					2820
	agaatcaact					2880
	attgtgcaaa					2940
cataataat	aagaattaac	tcactattca	atagatattt	cctataaaaa	addagadage	3000
	tgcagttgca					3060
	gageteteae					3120
teteetact	aagaaatcag	ttattaaatt	atcettatae	agtataggga	gaaccccacc	3180
calgetacti	ccaataaaat	ctaacttaat	tacccagata	tttttacata	tttagcaaca	3240
araaaaaret+	atcatttgac	tcaacttaat	toctttctct	ttcttttcat	ttcctaccta	3300
	tttttatttg					3360
	cagacaaggc					3420
	tgtattactt					3480
	attaaattta					3532
cccaaayaa	accuuactea	ccccycca				2422

<210> 8

<211> 475

<212> PRT

<213> Mus Musculus

<400> 8

-7-Ile Met Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile 10 Leu Ser His Ile Val Ser Leu Phe Cys Cys Val Cys Lys Arg Arg Lys 20 25 30 Lys Asp Lys Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp 40 Gln Lys Lys Leu His Asp Phe Glu Glu Glu Cys Val Glu Met Tyr Phe 50 55 60 Asp Glu Lys Asp Asp Lys Phe Asn Ser Gly Ser Glu Glu Arg Ile Arg 65 70 75 80 70 75 Val Thr Phe Glu Arg Val Glu Gln Met Ser Ile Gln Ile Lys Glu Val 85 90 Gly Asp Arg Val Asn Tyr Ile Lys Arg Ser Leu Gln Ser Leu Asp Ser 100 105 Gln Ile Gly His Leu Gln Asp Leu Ser Ala Leu Thr Val Asp Thr Leu 115 120 Lys. Thr Leu Thr Ala Gln Lys Ala Ser Glu Ala Ser Lys Val His Asn 135 140 Glu Ile Thr Arg Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn Leu 150 155 Ile Asp Asp Val Pro Val Arg Pro Leu Trp Glu Glu Pro Ser Ala Val 170 Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Arg Glu Ser Asn Asn 185 Pro Phe Leu Cys Asn Ile Phe Met Lys Asp Glu Lys Asp Pro Gln Tyr 200 Asn Leu Phe Gly Gln Asp Leu Pro Val Ile Pro Gln Arg Lys Glu Phe 215 Asn Ile Pro Glu Ala Gly Ser Ser Cys Gly Ala Leu Phe Pro Ser Ala 230 235 Val Ser Pro Pro Glu Leu Arg Gln Arg Arg His Gly Val Glu Met Leu 250 245 Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Pro Asn Ser Ser 265 Pro His Met Ser Ser Pro Pro Thr Lys Phe Ser Val Ser Thr Pro Ser 280 Gln Pro Ser Cys Lys Ser His Leu Glu Ser Thr Thr Lys Asp Gln Glu 295 Pro Ile Phe Tyr Lys Ala Ala Glu Gly Asp Asn Ile Glu Phe Gly Ala 310 315 Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr 330 325 Ser Asn Lys Ile Arg Glu Leu Leu Ser Asn Asp Thr Pro Glu Asn Thr 340 345 350 Leu Lys His Val Gly Ala Ala Gly Tyr Ser Glu Cys Cys Lys Thr Ser 360 Thr Ser Leu His Ser Val Gln Ala Glu Ser Cys Ser Arg Arg Ala Ser 375 380 Thr Glu Asp Ser Pro Glu Val Asp Ser Lys Ala Ala Leu Leu Pro Asp 390 395 Trp Leu Arg Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Gly Gly
405 410 415 405 410 Thr Leu Asn Gly Leu Ala Ser Pro Phe Lys Pro Val Leu Asp Thr Asn 425 420 430 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln 440 445 435 Ser Ile Pro Phe Val Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val 455 460 Tyr Pro Ser Gly Gly Arg Val Leu Pro Val Tyr 470

```
<210> 9
<211> 5433
<212> DNA
<213> Mus Musculus
<220>
<221> unsure
<222> (5094)...(5094)
<223> unknown
```

<400> 9

qqctqaaaqa qcctqaqctq tqcctctcca ttccactgct gtggcagggt cagaaatctt 60 120 ggatagagaa aaccttttgc aaacgggaat gtatctttgt aattcctagc acgaaagact ctaacaggtg ttgctgtggc cagttcacca accagcatat ccccctctg ccaagtgcaa 180 240 cacccagcaa aaatgaagag gaaagcaaac aggtggagac tcagcctgag aaatggtctg 300 ttgccaagca cacccagagc tacccaacag attcctatgg agttcttgaa ttccagggtg 360 goggatatto caataaagco atgtatatoo gtgtatoota tgacaccaag ccagactcac tgctccatct catggtgaaa gattggcagc tggaactccc caagctctta atatctgtgc 420 480 atggaggcct ccagaacttt gagatgcagc ccaagctgaa acaagtcttt gggaaaggcc tgatcaaggc tgctatgacc accggggcct ggatcttcac cgggggtgtc agcacaggtg 540 600 ttatcagcca cgtaggggat gccttgaaag accactcctc caagtccaga ggccgggttt gtgctatagg aattgctcca tggggcatcg tggagaataa ggaagacctg gttggaaagg 660 720 atgtaacaag agtgtaccag accatgtcca accetctaag taagetetet gtgctcaaca acteccacae ccaetteate etggetgaca atggeacet gggeaagtat ggegeegagg 780 840 tqaaqctqcq aaqgctqctq qaaaaqcaca tctccctcca gaagatcaac acaagactgg ggcagggcgt gcccctcgtg ggtctcgtgg tggagggggg ccctaacgtg gtgtccatcg 900 960 tcttggaata cctgcaagaa gagcctccca tccctgtggt gatttgtgat ggcagcggac 1020 gtgcctcgga catcctgtcc tttgcgcaca agtactgtga agaaggcgga ataataaatg 1080 agteceteag ggageagett etagttaeea tteagaaaae atttaattat aataaggeae 1140 aatcacatca gctgtttgca attataatgg agtgcatgaa gaagaaagaa ctcgtcactg 1200 tgttcagaat gggttctgag ggccagcagg acatcgagat ggcaatttta actgccctgc 1260 tgaaaggaac aaacgtatct gctccagatc agctgagctt ggcactggct tggaaccgcg tggacatage acgaagecag atctttgtct ttgggcccca ctggacgccc ctgggaagec 1320 tggcaccccc gacggacagc aaagccacgg agaaggagaa gaagccaccc atggccacca 1380 1440 ccaagggagg aagaggaaaa gggaaaggca agaagaaagg gaaagtgaaa gaggaagtgg 1500 aggaagaaac tgacccccgg aagatagagc tgctgaactg ggtgaatgct ttggagcaag 1560 eqatqctaqa tqctttaqtc ttagatcgtg tcgactttgt gaagctcctg attgaaaacg 1620 gagtgaacat gcaacacttt ctgaccattc cgaggctgga ggagctctat aacacaagac tgggtccacc aaacacactt catctgctgg tgagggatgt gaaaaagagc aaccttccgc 1680 1740 ctgattacca catcagcctc atagacatcg ggctcgtgct ggagtacctc atgggaggag cctaccgctg caactacact cggaaaaact ttcggaccct ttacaacaac ttgtttggac 1800 1860 caaagaggcc taaagctctt aaacttctgg gaatggaaga tgatgagcct ccagctaaag qqaaqaaaa aaaaaaaaq aaaaaggagg aagagatcga cattgatgtg gacgaccctg 1920 1980 ccqtqaqtcq qttccagtat cccttccacg agctgatggt gtgggcagtg ctgatgaaac gccagaaaat ggcagtgttc ctctggcagc gaggggaaga gagcatggcc aaggccctgg 2040 2100 tggcctgcaa gctctacaag gccatggccc acgagtcctc cgagagtgat ctggtggatg acatetecea ggaettggat aacaatteca aagaettegg ecagettget ttggagttat 2160 2220 tagaccagte ctataagcat gacgagcaga tegetatgaa acteetgace tacgagetga aaaactggag caactcgacc tgcctcaaac tggccgtggc agccaaacac cgggacttca 2280 2340 ttgctcacac ctgcagccag atgctgctga ccgatatgtg gatgggaaga ctgcggatgc ggaagaaccc cggcctgaag gttatcatgg ggattcttct accccccacc atcttgtttt 2400 2460 tggaatttcg cacatatgat gatttctcgt atcaaacatc caaggaaaac gaggatggca aagaaaaaga agaggaaaat acggatgcaa atgcagatgc tggctcaaga aagggggatg 2520 2580 aggagaacga gcataaaaaa cagagaagta ttcccatcgg aacaaagatc tgtgaattct 2640 ataacgcgcc cattgtcaag ttctggtttt acacaatatc atacttgggc tacctgctgc tgtttaacta cgtcatcctg gtgcggatgg atggctggcc gtccctccag gagtggatcg 2700 2760 tcatctccta catcgtgagc ctggcgttag agaagatacg agagatcctc atgtcagaac caggcaaact cagccagaaa atcaaagttt ggcttcagga gtactggaac atcacagatc 2820 2880 tegtggccat ttecacatte atgattggag caattetteg cetacagaae cageeetaca

```
tgggctatgg ccgggtgatc tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg
                                                                        2940
acatettigg tgteaacaag tatetgggge cataegtgat gatgattgga aagatgatga
                                                                        3000
togacatgot glacttigtg gloatcatgo tggtogtgot catgagtito ggagtageco
                                                                        3060
                                                                        3120
qtcaaqccat tctqcatcca qaqqaqaaqc cctcttggaa actggcccga aacatcttct
                                                                       3180
acatgcccta ctggatgatc tatggagagg tgtttgcaga ccagatagac ctctacgcca
 tggaaattaa teeteettgt ggtgagaace tatatgatga ggagggcaag eggetteete
                                                                        3240
                                                                        3300
 cctgtatccc cggcgcctgg ctcactccag cactcatggc gtgctatcta ctggtcgcca
acatectget ggtgaacetg etgattgetg tgttcaacaa tactttettt gaagtaaaat
                                                                        3360
caatatccaa ccaggtgtgg aagttccagc gatatcagct gattatgaca tttcatgaca
                                                                        3420
 ggccagtcct gccccaccg atgatcattt taagccacat ctacatcatc attatgcgtc
                                                                        3480
 tcagcggccg ctgcaggaaa aagagagaag gggaccaaga ggaacgggat cgtggattga
                                                                        3540
                                                                        3600
agetetteet tagegaegag gagetaaaga ggetgeatga gttegaggag eagtgegtge
aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcatccggg
                                                                        3660
                                                                        3720
tcacttctga aagagttgaa aatatgtcaa tgaggttgga agaaatcaat gaaagagaaa
cttttatgaa aacttccctg cagactgttg accttcgact tgctcagcta gaagaattat
                                                                        3780
ctaacagaat ggtgaatgct cttgaaaatc ttgcgggaat cgacaggtct gacctgatcc
                                                                        3840
                                                                        3900
aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccgg caaagcagca
                                                                        3960
tcaatagege tgatggetae agettgtate gatateattt taaeggagaa gagttattat
                                                                        4020
ttqaqqatac atctctctcc acgtcaccag ggacaggagt caggaaaaaa acctgttcct
                                                                        4080
tccgtataaa ggaagagaag gacgtgaaaa cgcacctagt cccagaatgt cagaacagtc
ttcacctttc actgggcaca agcacatcag caaccccaga tggcagtcac cttgcagtag
                                                                        4140
atgacttaaa gaacgctgaa gagtcaaaat taggtccaga tattgggatt tcaaaggaag
                                                                        4200
                                                                       4260
atgatgaaag acagacagac totaaaaaaag aagaaactat ttooccaagt ttaaataaaa
cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa
                                                                       4320
cgacaaatat agaaggcact atttcctatc ccctggaaga aaccaaaatt acacgctatt
                                                                       4380
teccegatga aacgateaat gettgtaaaa caatgaagte cagaagette gtetatteee
                                                                       4440
                                                                       4500
ggggaagaaa gctggtcggt ggggttaacc aggatgtaga gtacagttca atcacggacc
agcaattgac gacggaatgg caatgccaag ttcaaaagat cacgcgctct catagcacag
                                                                       4560
atatteetta cattgtgteg gaagetgeag tgeaagetga geaaaaagag cagtttgeag
                                                                        4620
                                                                       4680
atatgcaaga tgaacaccat gtcgctgaag caatteeteg aateeetege ttgteeetaa
ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg gattccatc tctcaggtca aaaagtttac atggacatcc taggaatgtg aaatccattc
                                                                        4740
                                                                       4800
                                                                        4860
agggaaagtt agacagatct ggacatgcca gtagtgtaag cagcttagta attgtgtctg
                                                                       4920
gaatgacagc agaagaaaaa aaggttaaga aagagaaagc ttccacagaa actgaatgct
                                                                       4980
agtotgttit gittotttaa tittititti taacagtoag aaacccacta aigggigtoa
tottggccca toctaaacac atmtccaatt toctaaaaac attttccctt aaaaaatttt
                                                                       5040
                                                                       5100
ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatatgt
                                                                       5160
tagtaggett agttttttca gttgcagtag tatcaaatga aagtgatgat actgtaacga
                                                                       5220
agataaattg gctaatcagt atacaagatt atacaatctc tttattactg agggccacca
                                                                       5280
aataqcctaq qaaqtqccct cqagcactga agtcaccatt aggtcactca agaagtaagc
aactagetgg geacagtgge teatgeetgt aateetagea etttgggagg ceaaggeaga
                                                                       5340
                                                                       5400
aagatagett gagteeagga gtttgagaee ageetgggea acatagtgat accecatete
ttaaaaaaa aaaaaaaaa ctgccctcgt gcc
                                                                       5433
```

<210> 10

<211> 1533

<212> PRT

<213> Mus Musculus

<400> 10

 Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His

 1
 5

 Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser

 20
 25

 Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln

 35
 40

 Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp

WO	VU/41	J014													101/
									-10-						
65					70				Val	75					80
				85					Arg 90					95	
-			100					105	Asn				110		
•	_	115					120		Met			125			•
	130					135			His		140				
145					150				Val	155					160
	_			165					Asn 170					175	
			180					185	Gly				190		
		195					200		Pro			205			
	210					215			Ile		220				
225					230				Glu Tyr	235					240
				245					250 Met					255	
			260					265	Gln				270		
		275					280		Asn			285			
	290					295			Val		300				
305					310				Pro	315					320
				325					330 Glu					335	
			340					345	Lys				350		
		355					360		Asp			365			
Leu	370 Asn	Trp	Val	Asn	Ala	375 Leu	Glu	Gln	Ala		380 Leu	Asp	Ala	Leu	
385 Leu	Asp	Arg	Val	Asp	390 Phe	Val	Lys	Leu	Leu	395 Ile	Glu	Asn	Gly	Val	400 Asn
Met	Gln	His		405 Leu	Thr	Ile	Pro		410 Leu	Glu	Glu	Leu		415 Asn	Thr
Arg	Leu		420 Pro	Pro	Asn	Thr		425 His	Leu	Leu	Val		430 Asp	Val	Lys
Lys		435 Asn	Leu	Pro	Pro		440 Tyr	His	Ile	Ser		445 Ile	Asp	Ile	Gly
	450 Val	Leu	Glu	Tyr		455 Met	Gly	Gly	Ala	Tyr 475	460 Arg	Суз	Asn	Tyr	Thr 480
465 Arg	Lys	Asn	Phe		470 Thr	Leu	Tyr	Asn	Asn 490		Phe	Gly	Pro	Lys 495	
Pro	Lys	Ala		485 Lys	Leu	Leu	Gly	Met 505	Glu	Asp	Asp	Glu	Pro 510		Ala
Lys	Gly	Lys 515	500 Lys	Lys	Lys	Lys	Lys 520		Lys	Glu	Glu	Glu 525		Asp	Ile
Asp	Val 530		Asp	Pro	Ala	Val 535		Arg	Phe	Gln	Tyr 540		Phe	His	Glu

-11-Leu Met Val Trp Ala Val Leu Met Lys Arg Gln Lys Met Ala Val Phe Leu Trp Gln Arg Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys Leu Tyr Lys Ala Met Ala His Glu Ser Ser Glu Ser Asp Leu Val Asp Asp Ile Ser Gln Asp Leu Asp Asn Asn Ser Lys Asp Phe Gly Gln Leu Ala Leu Glu Leu Leu Asp Gln Ser Tyr Lys His Asp Glu Gln Ile 610 615 620 Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys Leu Lys Leu Ala Val Ala Ala Lys His Arg Asp Phe Ile Ala His Thr Cys Ser Gln Met Leu Leu Thr Asp Met Trp Met Gly Arg Leu Arg Met Arg Lys Asn Pro Gly Leu Lys Val Ile Met Gly Ile Leu Leu Pro Pro Thr Ile Leu Phe Leu Glu Phe Arg Thr Tyr Asp Asp Phe Ser Tyr Gln Thr Ser Lys Glu Asn Glu Asp Gly Lys Glu Lys Glu Glu Glu Asn Thr Asp Ala Asn Ala Asp Ala Gly Ser Arg Lys Gly Asp Glu Glu Asn Glu His Lys Lys Gln Arg Ser Ile Pro Ile Gly Thr Lys Ile Cys Glu Phe Tyr Asn Ala Pro Ile Val Lys Phe Trp Phe Tyr Thr Ile Ser Tyr Leu Gly Tyr Leu Leu Phe Asn Tyr Val Ile Leu Val Arg Met Asp Gly Trp Pro Ser Leu Gln Glu Trp Ile Val Ile Ser Tyr Ile Val Ser Leu Ala Leu Glu Lys Ile Arg Glu Ile Leu Met Ser Glu Pro Gly Lys Leu Ser Gln Lys Ile Lys Val Trp Leu Gln Glu Tyr Trp Asn Ile Thr Asp Leu Val Ala Ile Ser Thr Phe Met Ile Gly Ala Ile Leu Arg Leu Gln Asn Gln Pro Tyr Met Gly Tyr Gly Arg Val Ile Tyr Cys Val Asp Ile Ile Phe Trp Tyr Ile Arg Val Leu Asp Ile Phe Gly Val Asn Lys Tyr Leu Gly Pro Tyr Val Met Met Ile Gly Lys Met Met Ile Asp Met Leu Tyr Phe Val Val Ile Met Leu Val Val Leu Met Ser Phe Gly Val Ala Arg Gln Ala Ile Leu His Pro Glu Glu Lys Pro Ser Trp Lys Leu Ala Arg Asn Ile Phe Tyr Met Pro Tyr Trp Met Ile Tyr Gly Glu Val Phe Ala Asp Gln Ile Asp Leu Tyr Ala Met Glu Ile Asn Pro Pro Cys Gly Glu Asn Leu Tyr Asp Glu Glu Gly Lys Arg Leu Pro Pro Cys Ile Pro Gly Ala Trp Leu Thr Pro Ala Leu Met Ala Cys Tyr Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Val Phe Asn Asn Thr 995 1000 Phe Phe Glu Val Lys Ser Ile Ser Asn Gln Val Trp Lys Phe Gln Arg

Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro 1030 1035 Met Ile Ile Leu Ser His Ile Tyr Ile Ile Ile Met Arg Leu Ser Gly 1045 1055 1050 Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly 1060 1065 1070 Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe 1075 1080 1085 Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln 1090 1095 1100 Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu 1110 1115 112 Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met 1125 1130 1135 Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu 1140 1145 1150 Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp 1160 1165 Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu 1175 1180 Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr 1190 1195 120 Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp 1205 1210 1215 Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys 1220 1225 1230 1220 Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro 1240 1245 Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala 1255 1260 1250 Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu 1270 1275 Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu 1285 1290 1295 Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn 1305 1310 1300 Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr 1320 1325 1315 Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro 1330 1335 1340 Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn 1345 1350 1355 136 Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg 1365 1370 Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr 1380 1385 1390 Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr 1395 1400 1405 Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val 1410 1415 1420 Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His 1430 1435 1425 Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr 1445 1450 1455 Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr 1460 1465 1470 Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg 1480 1485 1475 Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser

1495

1500

-13-

 Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys

 1505
 1510
 1515
 152

 Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
 1525
 1530

<210> 11 <211> 6220 <212> DNA <213> Homo Sapiens

<400> 11

tgtgcagaat tgtacagttg cgaaaccatg tcgctggcag ctggtgctgg cggtggagac 60 ttccctgtgc ggtgctcagt gcatctgcac ccgtggggga gggagctctt tctctggccc 120 tgcagtcacc tgaggttgtt accattatga acggccgctg ggacccccgc atgtgcatgt 180 240 actocccag agtgtccggg ggccccagcc aagggacaca tctcacgcag ctgggaacat 300 gtgcaggctg atgaagagaa ccggatgagg gcttcacatg aggaagcatg tggccaggtc 360 ctctcagaac atcagcctca tcttcctgtc tctgatctat ttcaccaacc accccatgtg tetetagaac eccagtgtag egagetggag agaggaetgt eetgagggea geaggeetgg 420 ttgcagctgg cgtgggggtc tcagaatgga gccctcagcc ctgaggaaag ctggctcgga 480 540 gcaggaggag ggctttgagg ggctgcccag aagggtcact gacctgggga tggtctccaa tctccggcgc agcaacagca gcctcttcaa gagctggagg ctacagtgcc ccttcggcaa 600 660 caatgacaag caagaaagcc tcagttcgtg gattcctgaa aacatcaaga agaaagaatg cgtgtatttt gtggaaagtt ccaaactgtc tgatgctggg aaggtggtgt gtcagtgtgg 720 780 ctacacgcat gagcagcact tggaggaggc taccaagccc cacaccttcc agggcacaca 840 gtgggaccca aagaaacatg tccaggagat gccaaccgat gcctttggcg acatcgtctt cacgggcctg agccagaagg tgaaaaagta cgtccgagtc tcccaggaca cgccctccag 900 960 cgtgatctac cacctcatga cccagcactg ggggctggac gtccccaatc tcttgatctc 1020 ggtgaccggg ggggccaaga acttcaacat gaagccgcgg ctgaagagca ttttccgcag 1080 aggectggte aaggtggete agaceaeagg ggeetggate ateaeagggg ggteeeaeae cggcgtcatg aagcaggtag gcgaggcggt gcgggacttc agcctgagca gcagctacaa 1140 ggaaggcgag ctcatcacca tcggagtcgc cacctggggc actgtccacc gccgcgaggg 1200 1260 cctgatccat cccacgggca gcttccccgc cgagtacata ctggatgagg atggccaagg gaacctgacc tgcctagaca gcaaccactc tcacttcatc ctcgtggacg acgggaccca 1320 1380 cggccagtac ggggtggaga ttcctctgag gaccaggctg gagaagttca tatcggagca 1440 gaccaaggaa agaggaggtg tggccatcaa gatccccatc gtgtgcgtgg tgctggaggg 1500 cggcccgggc acgttgcaca ccatcgacaa cgccaccacc aacggcaccc cctgtgtggt tgtggagggc tcgggccgcg tggccgacgt cattgcccag gtggccaacc tgcctgtctc 1560 ggacatcact atctccctga tccagcagaa actgagcgtg ttcttccagg agatgtttga 1620 gacetteacg gaaagcagga ttgtcgagtg gaccaaaaag atccaagata ttgtccggag 1680 gcggcagctg ctgactgtct tccgggaagg caaggatggt cagcaggacg tggatgtggc 1740 catcttgcag gccttgctga aagcctcacg gagccaagac cactttggcc acgagaactg 1800 1860 ggaccaccag ctgaaactgg cagtggcatg gaatcgcgtg gacattgccc gcagtgagat cttcatggat gagtggcagt ggaagcette agatetgcae cecaegatga cagetgeact 1920 1980 catctccaac aagcctgagt ttgtgaagct cttcctggaa aacggggtgc agctgaagga gtttgtcacc tgggacacct tgctctacct gtacgagaac ctggacccct cctgcctgtt 2040 ccacagcaag ctgcaaaagg tgctggtgga ggatcccgag cgcccggctt gcgcgcccgc 2100 ggcgccccgc ctgcagatgc accacgtggc ccaggtgctg cgggagctgc tgggggactt 2160 2220 cacgcagccg ctttatcccc ggccccggca caacgaccgg ctgcggctcc tgctgcccgt tececacgte aageteaacg tgeagggagt gageeteegg tecetetaca agegtteete 2280 2340 aggccatgtg accttcacca tggaccccat ccgtgacctt ctcatttggg ccattgtcca gaaccgtcgg gagctggcag gaatcatctg ggctcagagc caggactgca tcgcagcggc 2400 cttggcctgc agcaagatcc tgaaggaact gtccaaggag gaggaggaca cggacagctc 2460 ggaggagatg ctggcgctgg cggaggagta tgagcacaga gccatcgggg tcttcaccga 2520 gtgctaccgg aaggacgaag agagagccca gaaactgctc acccgcgtgt ccgaggcctg 2580 2640 ggggaagacc acctgcctgc agctcgccct ggaggccaag gacatgaagt ttgtgtctca cgggggcatc caggccttcc tgaccaaggt gtggtggggc cagetctccg tggacaatgg 2700 2760 getgtggcgt gtgaccetgt geatgetgge ettecegetg etceteaceg geeteatete cttcagggag aagaggctgc aggatgtggg cacccccgcg gcccgcgccc gtgccttctt 2820 2880 caccgcaccc gtggtggtct tccacctgaa catcctctcc tacttcgcct tcctctgcct gttcgcctac gtgctcatgg tggacttcca gcctgtgccc tcctggtgcg agtgtgccat 2940

			-1-4			
ctacctctgg	ctcttctcct	tggtgtgcga	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gracttcagt	gacttctgga	ataagctgga	3060
cgtcggcgca	atcttgctct	tcgtggcagg	gctgacctgc	aggctcatcc	cggcgacgct	3120
qtaccccqqq	cgcgtcatcc	tctctctgga	cttcatcctg	ttctgcctcc	ggctcatgca	3180
catttttacc	atcagtaaga	cgctggggcc	caagatcatc	attgtgaagc	ggatgatgaa	3240
ggacgtcttc	ttcttcctct	tcctgctggc	tgtgtgggtg	gtgtccttcg	gggtggccaa	3300
gcaggccatc	ctcatccaca	acgagcgccg	ggtggactgg	ctgttccgag	gggccgtcta	3360
ccactcctac	ctcaccatct	tcgggcagat	cccgggctac	atcgacggtg	tgaacttcaa	3420
cccqqaqcac	tgcagcccca	atggcaccga	cccctacaag	cctaagtgcc	ccgagagcga	3480
cqcqacqcaq	cagaggccgg	ccttccctga	gtggctgacg	gtcctcctac	tctgcctcta	3540
cctqctcttc	accaacatcc	tgctgctcaa	cctcctcatc	gccatgttca	actacacctt	3600
ccagcaggtg	caggagcaca	cggaccagat	ttggaagttc	cagcgccatg	acctgatcga	3660
ggagtaccac	ggccgccccg	ccgcgccgcc	ccccttcatc	ctcctcagcc	acctgcagct	3720
cttcatcaag	agggtggtcc	tgaagactcc	ggccaagagg	cacaagcagc	tcaagaacaa	3780
gctggagaag	aacgaggagg	cggccctgct	atcctgggag	atctacctga	aggagaacta	3840
cctccagaac	cgacagttcc	agcaaaagca	gcggcccgag	cagaagatcg	aggacatcag	3900
caataaggtt	gacgccatgg	tggacctgct	ggacctggac	ccactgaaga	ggtcgggctc	3960
catggagcag	aggttggcct	ccctggagga	gcaggtggcc	cagacagccc	gagccctgca	4020
ctggatcgtg	aggacgctgc	gggccagcgg	cttcagctcg	gaggcggacg	tececaetet	4080
ggcctcccag	aaggccgcgg	aggagccgga	tgctgagccg	ggaggcagga	agaagacgga	4140
ggagccgggc	gacagctacc	acgtgaatgc	ccggcacctc	ctctacccca	actgccctgt	4200
cacgcgcttc	cccgtgccca	acgagaaggt	gccctgggag	acggagttcc	tgatctatga	4260
cccacccttt	tacacggcag	agaggaagga	cacaaccacc	atggacccca	tgggagacac	4320
cctggagcca	ctgtccacga	tccagtacaa	cgtggtggat	ggcctgaggg	accgccggag	4380
cttccacggg	ccgtacacag	tgcaggccgg	gttgcccctg	aaccccatgg	gccgcacagg	4440
actgcgtggg	cgcgggagcc	tcagctgctt	cggacccaac	cacacgctgt	accccatggt	4500
cacgcggtgg	aggcggaacg	aggatggagc	catctgcagg	aagagcataa	agaagatgct	4560 4620
ggaagtgctg	gtggtgaagc	tccctctctc	cgagcactgg	gccctgcctg	ggggctcccg	4680
ggagccaggg	gagatgctac	ctcggaagct	gaagcggatc	ctccggcagg	ageaetggee	4740
gtcttttgaa	aacttgctga	agtgcggcat	ggaggtgtac	aaaggctaca	tggatgaccc	4800
gaggaacacg	gacaatgcct	ggatcgagac	ggtggccgtc	agegreeaer	caaaaaacta	4860
gaatgacgtg	gagctgaaca	ggctgaactc	caacetgeac	geetgegaet	agagggcccc	4920
catccgatgg	caggtggtgg	acaggegeat	cccactctat	tagastanaa	staggegget	4980
ccagaaggca	gccgctgagt	teggggetea	ccactgactg	stactages	taaccaaac	5040
ccagtccata	gacgttcccc	ccagaaacca	gggettetet	gaccagtgc	ccctcacqqc	5100
teaggetgtt	cctgggccct	tancetarta	gggtttggtg	gatcaagata	accordage	5160
tgccgcaagt	ctgctgcaga	agatagtera	aactggaagg	ccctaggeg	ccttagccat	5220
agageteaag	acagggcaca gctgggcctg	tacaaataaa	cocttagggg	gagtccactc	ccttcctaac	5280
tatatasaa	cgagcagctc	atccaccato	gaggtcattg	gagteeacce	agttccccgg	5340
agagtagga	tcccctgtgg	cccctcagg	cctatatcta	taaaaaaaaa	accetaceae	5400
teteeceaaa	agggcctcca	tatttcaaga	tacctcaaca	tagageetta	cctaacctaa	5460
actaggggg	ctgtctgaac	tectaactat	caggataaac	tecatagaa	tacaggagcc	5520
cadacaaadc	ccaggcctgt	caagagacgc	agagggcccc	taccagggtt	ggccccaggg	5580
accetagas	gaggctgcag	aageteteee	tecetactee	ctgggagcca	catactaacc	5640
acceegggae	ggacggcatg	addetecece	aggacatag	agacetteta	atttaatatc	5700
acgeggeeag	aggagcgtga	accatgaggg	ccctcaggag	gggaacgtgg	taaaacccaa	5760
gacattagat	ctgccatctc	aggetaget	aactetteta	toctttccac	aaataaagtt	5820
cctgacacgt	ccagggccag	agactatata	acggctgcct	gaagttctcc	tcgatccccc	5880
agtgacatge	ctgcagcctg	tagatatact	gcagcccctc	agccctaccc	ccaagtttct	5940
cototoacco	atcagctccc	totcttcatt	ttcctaaacc	tgggctccag	categteece	6000
aagcccacca	ggccaggatg	caggcatcca	catqccctcc	tecttagett	cccctgcgtg	6060
gtggtgccaa	tgtgccctgg	cacccctgca	gaggeteegg	atggagcctg	gggctgcctg	6120
gccactgage	actggccgag	gtgatgccca	cccttcccta	gacaggeete	tgtcttccac	6180
ctgacccaaa	gctctctagc	cacccctto	tccccagtat		-	6220
2-3-300000	, 		,			

<212> PRT <213> Homo Sapiens

<400> 12 Met Glu Pro Ser Ala Leu Arg Lys Ala Gly Ser Glu Gln Glu Gly 5 10 Phe Glu Gly Leu Pro Arg Arg Val Thr Asp Leu Gly Met Val Ser Asn 25 Leu Arg Arg Ser Asn Ser Ser Leu Phe Lys Ser Trp Arg Leu Gln Cys 40 4.5 Pro Phe Gly Asn Asn. Asp Lys Gln Glu Ser Leu Ser Ser Trp Ile Pro 55 Glu Asn Ile Lys Lys Lys Glu Cys Val Tyr Phe Val Glu Ser Ser Lys 70 75 Leu Ser Asp Ala Gly Lys Val Val Cys Gln Cys Gly Tyr Thr His Glu 90 85 Gln His Leu Glu Glu Ala Thr Lys Pro His Thr Phe Gln Gly Thr Gln 110 100 105 Trp Asp Pro Lys Lys His Val Gln Glu Met Pro Thr Asp Ala Phe Gly Asp Ile Val Phe Thr Gly Leu Ser Gln Lys Val Lys Lys Tyr. Val Arg 130 135 Val Ser Gln Asp Thr Pro Ser Ser Val Ile Tyr His Leu Met Thr Gln 155 150 His Trp Gly Leu Asp Val Pro Asn Leu Leu Ile Ser Val Thr Gly Gly 175 170 165 Ala Lys Asn Phe Asn Met Lys Pro Arg Leu Lys Ser Ile Phe Arg Arg 190 185 180 Gly Leu Val Lys Val Ala Gln Thr Thr Gly Ala Trp Ile Ile Thr Gly 205 200 Gly Ser His Thr Gly Val Met Lys Gln Val Gly Glu Ala Val Arg Asp 220 215 Phe Ser Leu Ser Ser Ser Tyr Lys Glu Gly Glu Leu Ile Thr Ile Gly 235 230 Val Ala Thr Trp Gly Thr Val His Arg Arg Glu Gly Leu Ile His Pro 245 250 255 Thr Gly Ser Phe Pro Ala Glu Tyr Ile Leu Asp Glu Asp Gly Gln Gly 260 265 Asn Leu Thr Cys Leu Asp Ser Asn His Ser His Phe Ile Leu Val Asp 280 275 Asp Gly Thr His Gly Gln Tyr Gly Val Glu Ile Pro Leu Arg Thr Arg 295 Leu Glu Lys Phe Ile Ser Glu Gln Thr Lys Glu Arg Gly Gly Val Ala 315 310 Ile Lys Ile Pro Ile Val Cys Val Val Leu Glu Gly Gly Pro Gly Thr 325 330 Leu His Thr Ile Asp Asn Ala Thr Thr Asn Gly Thr Pro Cys Val Val 345 340 Val Glu Gly Ser Gly Arg Val Ala Asp Val Ile Ala Gln Val Ala Asn 365 360 355 Leu Pro Val Ser Asp Ile Thr Ile Ser Leu Ile Gln Gln Lys Leu Ser 380 375 Val Phe Phe Gln Glu Met Phe Glu Thr Phe Thr Glu Ser Arg Ile Val 390 395 Glu Trp Thr Lys Lys Ile Gln Asp Ile Val Arg Arg Arg Gln Leu Leu 405 410 Thr Val Phe Arg Glu Gly Lys Asp Gly Gln Gln Asp Val Asp Val Ala 425 Ile Leu Gln Ala Leu Leu Lys Ala Ser Arg Ser Gln Asp His Phe Gly 440

WC	00/4	0614													PCI
								,	-16-						
His	Glu 450	Asn	Trp	Asp	His	Gln 455	Leu	Lys	Leu	Ala	Val 460	Ala	Trp	Asn	Àrg
Val 465	Asp	Ile	Ala	Arg	Ser 470		Ile	Phe	Met	Asp 475	Glu	Trp	Gln	Trp	Lys 480
	Ser	Asp	Leu	His 485		Thr	Met	Thr	Ala 490	Ala	Leu	Ile	Ser	Asn 495	Lys
Pro	Glu	Phe	Val 500	Lys	Leu	Phe	Leu	Glu 505	Asn	Gly	Val	Gln	Leu 510	Lys	Glu
		515					520					525		Asp	
	530					535					540			Asp	
545					550					555				His	560
				565					570					Pro 575	
			580					585					590	Pro	
		595					600					605		Leu	
	610					615					620			Arg	
625					630					635				Gly	640
				645					650					Cys 655	
			660					665					670	Ser	
		675					680					685		Ile	
	690					695					700			Lys Gln	
705					710					715				Ile	720
				725					730					735 Asn	
			740					745					750	Leu	
		755					760					765		Thr	
_	770				-	775					780			Phe	
785					790					795				Tyr	800
				805					810					815 Ala	
			820					825					830	Leu	
		835					840					845		Tyr	
	850					855					860			Phe	
865					870					875				Gly	880
				885					890					895 Met	
			900					905					910	Val	
110	T 11E	915	116	JUL	- y 5		920	~ _ y	0	-,5		925			_,_

-17-

```
Arg Met Met Lys Asp Val Phe Phe Phe Leu Phe Leu Leu Ala Val Trp
                935
                                 940
Val Val Ser Phe Gly Val Ala Lys Gln Ala Ile Leu Ile His Asn Glu
               950
                     955
Arg Arg Val Asp Trp Leu Phe Arg Gly Ala Val Tyr His Ser Tyr Leu
                970 975
           965
Thr Ile Phe Gly Gin Ile Pro Gly Tyr Ile Asp Gly Val Asn Phe Asn 980 985 990
Pro Glu His Cys Ser Pro Asn Gly Thr Asp Pro Tyr Lys Pro Lys Cys
          1000 1005
Pro Glu Ser Asp Ala Thr Gln Gln Arg Pro Ala Phe Pro Glu Trp Leu
                1015
                                  1020
Thr Val Leu Leu Cys Leu Tyr Leu Leu Phe Thr Asn Ile Leu Leu
                      1035
             1030
Leu Asn Leu Leu Ile Ala Met Phe Asn Tyr Thr Phe Gln Gln Val Gln
                   1050
                                          1055
          1045
Glu His Thr Asp Gln Ile Trp Lys Phe Gln Arg His Asp Leu Ile Glu
                       1065
                                       1070
       1060
Glu Tyr His Gly Arg Pro Ala Ala Pro Pro Pro Phe Ile Leu Leu Ser
1075 1080 1085
     1075 1080
His Leu Gln Leu Phe Ile Lys Arg Val Val Leu Lys Thr Pro Ala Lys
 1090 1095 1100
Arg His Lys Gln Leu Lys Asn Lys Leu Glu Lys Asn Glu Glu Ala Ala
1105 1110
                              1115
Leu Leu Ser Trp Glu Ile Tyr Leu Lys Glu Asn Tyr Leu Gln Asn Arg
     1125 1130 1135
Gln Phe Gln Gln Lys Gln Arg Pro Glu Gln Lys Ile Glu Asp Ile Ser
                        1145 1150
        1140
Asn Lys Val Asp Ala Met Val Asp Leu Leu Asp Leu Asp Pro Leu Lys
 1155
                    1160
Arg Ser Gly Ser Met Glu Gln Arg Leu Ala Ser Leu Glu Glu Gln Val
 1170 1175
                                  1180
Ala Gln Thr Ala Arg Ala Leu His Trp Ile Val Arg Thr Leu Arg Ala
       1190 1195
Ser Gly Phe Ser Ser Glu Ala Asp Val Pro Thr Leu Ala Ser Gln Lys
          1205
                         1210
Ala Ala Glu Glu Pro Asp Ala Glu Pro Gly Gly Arg Lys Thr Glu
        1220 1225
Glu Pro Gly Asp Ser Tyr His Val Asn Ala Arg His Leu Leu Tyr Pro
     1235 1240
                                    1245
Asn Cys Pro Val Thr Arg Phe Pro Val Pro Asn Glu Lys Val Pro Trp
         1255 1260
 1250
Glu Thr Glu Phe Leu Ile Tyr Asp Pro Pro Phe Tyr Thr Ala Glu Arg
       1270 1275
1265
Lys Asp Ala Ala Ala Met Asp Pro Met Gly Asp Thr Leu Glu Pro Leu
            1285
                          1290 1295
Ser Thr Ile Gln Tyr Asn Val Val Asp Gly Leu Arg Asp Arg Arg Ser
                      1305
        1300
                             1310
Phe His Gly Pro Tyr Thr Val Gln Ala Gly Leu Pro Leu Asn Pro Met
               1320 1325
     1315
Gly Arg Thr Gly Leu Arg Gly Arg Gly Ser Leu Ser Cys Phe Gly Pro
        1335 1340
  1330
Asn His Thr Leu Tyr Pro Met Val Thr Arg Trp Arg Arg Asn Glu Asp
               1350 1355 136
Gly Ala Ile Cys Arg Lys Ser Ile Lys Lys Met Leu Glu Val Leu Val
                           1370 1375
           1365
Val Lys Leu Pro Leu Ser Glu His Trp Ala Leu Pro Gly Gly Ser Arg
        1380
                       1385
                                        1390
Glu Pro Gly Glu Met Leu Pro Arg Lys Leu Lys Arg Ile Leu Arg Gln
                     1400
                                     1405
```

PCT/US99/29996 WO 00/40614

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val 1420 1415 Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile 144 1435 1430 Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu 1445 1450 1455 Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser 1460 1465 1470 Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His 1485 1480 1475 Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr 1495

<210> 13 <211> 1816 <212> PRT <213> C. Elegans

<400> 13

Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn 10 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser 30 20 25 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr 45 40 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu 55 60 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser 70 75 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu 90 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg 110 105 100 Glu Thr Glu Thr Gln Thr Lys Arg Arg Arg Lys Lys Gln Arg Ser Thr 120 125 115 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys 135 140 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys 155 150 Arg Gly Asn Ser Thr Thr Ser His Ser Gly His Ala Thr Arg Ala Gly 165 170 175 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln 185 190 180 Cys Ser Tyr Phe Val.Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys 195 200 205 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln 215 220 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu 230 235 Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asp Glu Val Asn Leu Thr 245 250 255 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp 260 265 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr 285 275 280 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met 300 295 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn 310 315 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe -19-

				225					220					225	
_	_	- 1		325		T 3 -		T	330	m L	7	7.1.	m	335	T10
_	-		340					345					350	Ile	
		355					360					365		Ala	
His	Ala 370	Leu	Glu	Phe	Trp	Ser 375	Phe	Gly	Leu	Phe	Trp 380	Val	Ile	Gln	Leu
Asp 385	Val	Leu	Leu	Ala	His 390	Ser	Met	Phe	Ile	Pro 395	Arg	Gly	Ser	Leu	Phe 400
	His	Gly	Asn	His 405	Thr	Ser	Lys	Asn	His 410	Val	Val	Ala	Ile	Gly 415	Ile
Ala	Ser	Trp	Gly 420		Leu	Lys	Gln	Arg 425		Arg	Phe	Val	Gly 430	Lys	Asp
Ser	Thr	Val 435		Tyr	Ala	Thr	Asn 440	Val	Phe	Asn	Asn	Thr 445	Arg	Leu	Lys
Glu	Leu 450		Asp	Asn	His	Ser 455	Tyr	Phe	Leu	Phe	Ser 460	Asp	Asn	Gly	Thr
Val 465		Arg	Tyr	Gly	Ala 470		Ile	Ile	Met	Arg 475	Lys	Arg	Leu	Glu	Ala 480
Tyr	Leu	Ala	Gln	Gly 485		Lys	Lys	Arg	Ser 490		Ile	Pro	Leu	Val 495	Cys
Val	Val	Leu	Glu 500		Gly	Ala	Phe	Thr 505		Lys	Met	Val	His 510	Asp	Tyr
Val	Thr	Thr 515		Pro	Arg	Ile	Pro 520		Ile	Val	Суз	Asp 525	Gly	Ser	Gly
Arg	Ala 530		Asp	Ile	Leu	Ala 535		Ala	His	Gln	Ala 540	Val	Ser	Gln	Asn
Gly 545		Leu	Ser	Asp	Asn 550	Ile	Arg	Asn	Gln	Leu 555	Val	Asn	Ile	Val	Arg 560
Arg	Ile	Phe	Gly	Tyr 565	Asp	Pro	Lys	Thr	Ala 570	Gln	Lys	Leu	Ile	Lys 575	Gln
Ile	Val	Glu	Cys 580		Thr	Asn	Lys	Ser 585	Leu	Met	Thr	Ile	Phe 590	Arg	Leu
Gly	Glu	Ser 595		Arg	Glu	Asp	Leu 600	Asp	His	Val	Ile	Met 605	Ser	Cys	Leu
Leu	Lys 610	Gly	Gln	Asn	Leu	Ser 615	Pro	Pro	Glu	Gln	Leu 620	Gln	Leu	Ala	Leu
Ala 625	Trp	Asn	Arg	Ala	Asp 630	Ile	Ala	Arg	Thr	Glu 635	Ile	Phe	Ala	Asn	Gly 640
Thr	Glu	Trp	Thr	Thr 645	Gln	Asp	Leu	His	Asn 650	Ala	Met	Ile	Glu	Ala 655	Leu
Ser	Asn	Asp	Arg 660	Ile	Asp	Phe	Val	His 665	Leu	Leu	Leu	Glu	Asn 670	Gly	Val
Ser	Met	Gln 675	Lys	Phe	Leu	Thr	Tyr 680	Gly	Arg	Leu	Glu	His 685	Leu	Tyr	Asn
Thr	Asp 690	Lys	Gly	Pro	Gln	Asn 695	Thr	Leu	Arg	Thr	Asn 700	Leu	Leu	Val	Asp
Ser 705	Lys	His	His	Ile	Lys 710	Leu	Val	Glu	Val	Gly 715	Arg	Leu	Val	Glu	Asn 720
Leu	Met	Gly	Asn	Leu 725	Tyr	Lys	Ser	Asn	Tyr 730	Thr	Lys	Glu	Glu	Phe 735	Lys
Asn	Gln	Tyr	Phe	Leu	Phe	Asn	Asn	Arg 745	Lys	Gln	Phe	Gly	Lys 750	Arg	Val
His	Ser	Asn 755	Ser	Asn	Gly	Gly	Arg 760	Asn	Asp	Val	Ile	Gly 765	Pro	Ser	Gly
Asp	Ala 770		Arg	Glu	Arg	Met 775		Ser	Met	Gln	Ile 780	Ser	Leu	Ile	Asn
Asn 785		Arg	Asn	Ser	Ile 790		Ser	Leu	Phe	Asn 795	Gly	Gly	Gly	Arg	Lys 800
	Glu	Ser	Asp	Asp		Asp	Asp	Phe	Ser	Asn	Leu	Glu	Glu	Glu	Ala

-20-

				805					810					815	
			820					825	Tyr				830		
		835					840		Ala			845			
	850					855			Val		860				
Ser 865	Leu	Ala	Lys	Thr	Ala 870	Ser	Leu	Ala	Thr	Gly 875	Glu	Ile	Gly	Met	Ser 880
Gln	Asp	Phe	Thr	Glu 885	Phe	Ser	Asp	Glu	Phe 890	Ser	Glu	Leu	Ala	Val 895	Glu
Val	Leu	Glu	Tyr 900	Cys	Thr	Lys	His	Gly 905	Arg	Asp	Gln	Thr	Leu 910	Arg	Leu
Leu	Thr	Cys 915	Glu	Leu	Ala	Asn	Trp 920	Gly	Asp	Glu	Thr	Cys 925	Leu	Ser	Leu
Ala	Ala 930	Asn	Asn	Gly	His	Arg 935	Lys	P.he	Leu	Ala	His 940	Pro	Cys	Cys	Gln
945					950				Gly	955					960
Gln	Asn	Ser	Lys	Val 965	Leu	Thr	Cys	Leu	Ala 970	Ala	Pro	Pro	Leu	Ile 975	Phe
Leu	Leu	Gly	Phe 980	Lys	Thr	Lys	Glu	Gln 985	Leu	Met	Leu	Gln	Pro 990	Lys	Thr
Ala	Ala	Glu 995	His	Asp	Glu	Glu	Met 1000		Asp	Ser	Glu	Met 1005		Ser	Ala
Glu	Asp 1010		Asp	Thr	Ser	Ser 1015		Ser	Ser	Ser	Asp 1020		Asp	Asp	Ser
Asp 1025		Glu	Asp	Ala	Lys 1030		Arg	Ala	Gln	Ser 103		Ser	Ala	Asp	Gln 104
Pro	Leu	Ser	Ile	His 1049		Leu	Val	Arg	Asp 1050		Leu	Asn	Phe	Ser 1055	
Lys	Lys	Lys	Pro 1060		Met	Gly	Ile	Ser 106	Arg 5	Ile	Val	Val	Ala 1070		Pro
Ile	Val	Thr 1075		Arg	Asn	Arg	Ala 1080		Thr	Met	Ser	Ile 1085		Lys	Ser
	1090)				1099	5		Cys		1100)			
Asp 1105		Asp	Glu	Gln	Glu 1110		Lys	Lys	Ala	Thr 1115		Met	Cys	Lys	Ser 112
Thr	Phe	Phe	Asp	Phe 1125		Phe	Asp	Phe	Pro 1130		Ile	Asn	Arg	Thr 1135	
Lys	Arg	Gly	Ser 1140		Ala	Val	Ala	Met 1145	Asn 5	His	Asp	Asp	Met 1150		Ile
_		1155	5				1160)	Thr			1165	5		
	1170)				1175	5		Val		1180)			
		Ser	Trp	Lys	Lys	Lys	Ile	Met	Glu	Phe 1195	Tyr	Lys	Ala	Pro	Ile 120
1185 Thr		Tyr	Trp	Leu 1205			Phe	Ala	Phe 1210	Ile		Phe	Leu	Ile 1215	Leu
Leu	Thr	Tyr	Asn 1220	Leu		Val	Lys	Thr 1225	Gln		Ile	Ala	Ser 1230	Trp	
Glu	Trp	Tyr 1235	Val		Ala	Tyr	Ile 1240	Phe	Val	Trp	Thr	Leu 1245	Glu		Gly
Arg	Lys 1250	Val		Ser		Ile 1255	Met		Asp	Thr	Ser 1260	Lys		Val	Leu
	Gln		Arg	Val	Phe	Phe		Gln	Tyr		Asn		Leu	Leu	
1265			_	 ,	1270		T 1 a	ת ז ת	m	1275		Δrσ	Lou	802	128 Pro
Phe	Glv	Leu	Leu	Tnr	Tvr	Leu	ı ı e	Ald	IVI	FILE	T T C	TT U	ne u	Set	110

	1285	7	1290		1295
Thr Thr Lys Thr	Leu Gly Arg 1		Ile Ile Cys	Asn Ser 1310	Val Ile
Trp Ser Leu Lys	Leu Val Asp 7	Tyr Leu S 1320	Ser Val Gln	1325	
Pro Tyr Ile Asn	1335		1340)	
Cys Val Leu Val	1350		1355		136
Ser Ile Thr Tyr	1365	:	1370		1375
Ile Phe Leu Gln 1380	Pro Tyr Phe N	1385		1390)
Glu Ile Asp Thr 1395		1400		1405	
Asn Ile Pro Ile 1410	1415		1420)	
Gly Tyr Trp Ile	1430		1435		144
Asn Val Leu Leu	1445		1450		1455
Glu Lys His Ile 1460)	1465		1470)
Gly Gln Val Met 1475		1480		1485	
Thr Ile Ile Tyr 1490	1495		150	0	
Ser Arg Met Phe 1505	1510		1515		152
Phe Leu Ser Pro	1525		1530		1535
Ser Val Glu Asp 1540	ו	1545		1550)
Asn Asp Glu Arg 1555		1560		1565	
Asn Arg Val Ser 1570	1575		158	0	
Ile Arg Glu Leu 1585	1590		1595		160
Glu Gln Met Asn	1605		1610		1615
Ile Ser Gly Lys 1620)	1625		163	U
Gly Gly Gly 1635		1640		1645	
Val Pro Met Ile 1650	1655		166	0	
Thr Ser Gly Gln 1665	1670		16/5		100
Ile Thr Glu Asn	1685		1690		1695
Ile Gln Phe Asn 1700	0	1705	•	171	0
Ala Thr Glu Glu 1715		1720		1725	
Arg Gln Val Thr 1730	1735		174	0	
Ile Thr Arg Glu	1750		1755		176
Pro Arg Arg His		o m1	TI- NI- NOS	Ala Tla	Clu Thr

PCT/US99/29996

WO 00/40614 -22-1770 1765 Glu Asp Asp Phe Tyr Ala Asp Ser Pro Val Pro Met Pro Met Thr Pro 1780 1785 1790 Val Gln Pro Ala Asp Gly Ser Phe Phe Gly Glu Asn Asp Ser Arg Tyr 1800 Gln Arg Asp Asp Ser Asp Tyr Glu 1815 <210> 14 <211> 1387 <212> PRT <213> C. Elegans <400> 14 Met Arg Lys Ser Arg Arg Val Arg Lys Leu Val Arg His Ala Ser Leu 10 Ile Glu Asn Ile Arg His Arg Thr Ser Ser Phe Leu Arg Leu Leu Asn 25 20 Ala Pro Arg Asn Ser Met Cys Asn Ala Asn Thr Val His Ser Ile Ser 35 40 Ser Phe Arg Ser Asp His Leu Ser Arg Lys Ser Thr His Lys Phe Leu 60 55 Asp Asn Pro Asn Leu Phe Ala Ile Glu Leu Thr Glu Lys Leu Ser Pro 75 70 Pro Trp Ile Glu Asn Thr Phe Glu Lys Arg Glu Cys Ile Arg Phe Ala 90 85 Ala Leu Pro Lys Asp Pro Glu Arg Cys Gly Cys Gly Arg Pro Leu Ser 105 100 Ala His Thr Pro Ala Ser Thr Phe Phe Ser Thr Leu Pro Val His Leu 125 120 Leu Glu Lys Glu Gln Gln Thr Trp Thr Ile Ala Asn Asn Thr Gln Thr 135 140 Ser Thr Thr Asp Ala Phe Gly Thr Ile Val Phe Gln Gly Gly Ala His 150 155 Ala His Lys Ala Gln Tyr Val Arg Leu Ser Tyr Asp Ser Glu Pro Leu 165 170 175 Asp Val Met Tyr Leu Met Glu Lys Val Trp Gly Leu Glu Ala Pro Arg 190 180 185 Leu Val Ile Thr Val His Gly Gly Met Ser Asn Phe Glu Leu Glu Glu 195 200 205 Arg Leu Gly Arg Leu Phe Arg Lys Gly Met Leu Lys Ala Ala Gln Thr 215 220 Thr Gly Ala Trp Ile Ile Thr Ser Gly Leu Asp Ser Gly Val Val Arg

235 230 His Val Ala Lys Ala Leu Asp Glu Ala Gly Ile Ser Ala Arg Met Arg 250 245 Ser Gln Ile Val Thr Ile Gly Ile Ala Pro Trp Gly Val Ile Lys Arg 265 270 260 Lys Glu Arg Leu Ile Arg Gln Asn Glu His Val Tyr Tyr Asp Val His 285 280 275 Ser Leu Ser Val Asn Ala Asn Val Gly Ile Leu Asn Asp Arg His Ser 300 295 Tyr Phe Leu Leu Ala Asp Asn Gly Thr Val Gly Arg Phe Gly Ala Asp 315 310 Leu His Leu Arg Gln Asn Leu Glu Asn His Ile Ala Thr Phe Gly Cys 325 335 330 Asn Gly Arg Lys Val Pro Val Val Cys Thr Leu Leu Glu Gly Gly Ile 340 345 350 Ser Ser Ile Asn Ala Ile His Asp Tyr Val Thr Met Lys Pro Asp Ile 360

Pro	Ala 370	Ile	Val	Cys	Asp	Gly 375	Ser	Gly	Arg	Ala	Ala 380	Asp	Ile	Ile	Ser
Phe 385	Ala	Ala	Arg	Tyr	Ile 390	Asn	Ser	Asp	Gly	Thr 395	Phe	Ala	Ala	Glu	Val 400
	Glu	Lys	Leu	Arg 405		Leu	Ile	Lys	Met 410		Phe	Pro	Glu	Thr 415	Asp
Gln	Glu	Glu	Met 420		Arg	Lys	Ile	Thr 425		Cys	Val	Ile	Arg 430	Asp	Asp
Leu	Leu	Arg		Phe	Arg	Tyr	Gly 440		Glu	Glu	Glu	Glu 445		Val	Asp
Phe	Val 450		Leu	Ser	Thr	Val 455		Gln	Lys	Gln	Asn 460		Pro	Pro	Asp
Glu 465		Leu	Ala	Leu	Thr 470		Ser	Trp	Asn	Arg 475		Asp	Leu	Ala	Lys 480
	Cys	Leu	Phe	Ser 485		Gly	Arg	Lys	Trp 490		Ser	Asp	Val	Leu 495	
Lys	Ala	Met	Asn 500		Ala	Leu	Tyr	Trp 505		Arg	Val	Asp	Phe 510	Val	Glu
Cys	Leu	Leu 515		Asn	Gly	Val	Ser 520	Met	Lys	Asn	Phe	Leu 525	Ser	Ile	Asn
Arg	Leu 530		Asn	Leu	Tyr	Asn 535		Asp	Asp	Ile	Asn 540	Ser	Ala	His	Ser
Val 545		Asn	Trp	Met	Glu 550	Asn	Phe	Asp	Ser	Met 555	Asp	Pro	His	Thr	Tyr 560
	Thr	Ile	Pro	Met 565		Gly	Gln	Val	Val 570	Glu	Lys	Leu	Met	Gly 575	Asn
Ala	Phe	Gln	Leu 580	Tyr	Tyr	Thr	Ser	Arg 585	Ser	Phe	Lys	Gly	Lys 590	Tyr	Asp
Arg	Tyr	Lys 595	Arg	Ile	Asn	Gln	Ser 600	Ser	Tyr	Phe	His	Arg 605	Lys	Arg	Lys
	610		_			615	_	_	_		620			Ile	
625					630					635				Asp	640
				645					650					Cys 655	
			660					665					670	Ile	
	_	675					680					685		Asp	
	690		_			695					700			Glu	
705					710		_	_	_	715				Arg	720
	_			72Š					730					Asn 735	
			740					745					750	His	
_	_	755					760					765		Lys	
	770					775					780			Pro	
785					790					795				Ala	800
				805					810					Glu 815	
			820					825					830	Pro	
Glu	Gln	Lys 835	Glu	Thr	Leu	Leu	Glu 840	Lys	Gly	Ser	Tyr	Thr 845	Lys	Lys	Val

Thr	Ile 850	Ile	Ser	Ser	Arg	Lys 855	Asn	Ser	Gly	Val	Ala 860	Ser	Val	Туг	Gly
	Ala	Ser	Ser	Met	Met 870		Lys	Arg	Glu	Pro 875	Gln	Leu	Asn	Lys	Phe 880
865	_		_			m	C ~ ~	C	Dro		The	Tue	Pho	Trn	
GLu	Arg	Pne	Arg		Pne	Tyr	Ser	ser	890	116	1111	Lys	rne	895	361
	_			885	_	-1 -	D	.		m	C1 -	mb	C		T ON
Trp	Cys	Ile		Phe	Leu	тте	Pne		Inr	Inr	GIN	Thr	Cys	TIE	Leu
			900					905		_		_	910		
Leu	Leu	Glu	Thr	Ser	Leu	Lys	Pro	Ser	Lys	Tyr	Glu	Trp	lle	Thr	Phe
		915					920					925			
Ile	Tyr	Thr	Val	Thr	Leu	Ser	Val	Glu	His	Ile	Arg	Lys	Leu	Met	Thr
	930					935					940				
Ser	Glu	Glv	Ser	Arg	Ile	Asn	Glu	Lys	Val	Lys	Val	Phe	Tyr	Ala	Lys
945					950					955					960
Trp	Tur	Asn	Ile	Trp	Thr	Ser	Ala	Ala	Leu	Leu	Phe	Phe	Leu	Val	Gly
	- , -			965					970					975	
Tur	Glv	Phe	Ara	Len	Val	Pro	Met	Tvr	Ara	His	Ser	Trp	Gly	Arq	Val
IYL	Gry	1116	980	Deu	•			985	9				990		
*	т	C	Dho	Co-	Acn	17 - 1	Lou		Tur	Met	Lvs	Ile		Glu	Tvr
Leu	Leu		rne	ser	ASII	vai	1000		1 y 1	1100	Lyo	1005	`		- 3 -
_	_	995		_	•	7			т	Tlo	C1 n			λla	Tue
Leu			HIS	Pro	Leu			PIO	IAT	TIE	9111	Met	АТА	VI G	БУЗ
	1010)			_	1015	٠.,	_	** . 7	•	1020		17-1	Dwa	т о
Met	Val	Trp	Ser	Met			тте	Cys	vaı	Leu	_Leu	Leu	vaı	PIO	Leu
1025	5				1030					103		_	_		104
Met	Ala	Phe	Gly	Val	Asn	Arg	Gln	Ala			Glu	Pro	Asn	Val	гуs
				1045					1050					1059	
Asp	Trp	His	Trp	Leu	Leu	Val	Arg	Asn	Ile	Phe	Tyr	Lys	Pro	Tyr	Phe
			1060)				106	5				1070)	
Met	Leu	Tvr	Glv	Glu	Val	Tyr	Ala	Gly	Glu	Ile	Asp	Thr	Cys	Gly	Asp
		1075	5				1080)				1085	5		
Glu	Gly	Tlo	70	0		_		m		T1 ~	D	D	~ .	-	
			ATO	CVS	Phe	Pro	GIV	TVI	Pne	тте	PFO	PIO	ьeu	Leu	met
Olu			Arg	Cys	Phe			Tyr	Pne	ire	110	0	Leu	Leu	мет
	1090)				1095	5				110	0			
Val	1090 Ile)			Val	1095 Ala	5			Leu	110 Leu	Asn			
Val	1090 Ile) Phe	Leu	Leu	Val	1095 Ala O	Asn	Ile	Leu	Leu 111	1100 Leu 5	0 Asn	Leu	Leu	Ile 112
Val	1090 Ile) Phe	Leu	Leu Asn	Val 1110 Ile	1095 Ala O	Asn	Ile	Leu Ser	Leu 111! Ile	1100 Leu 5	0	Leu	Leu Lys	Ile 112 Glu
Val 1109 Ala	1090 Ile 5 Ile	Phe Phe	Leu Asn	Leu Asn 1125	Val 1110 Ile	1099 Ala O Tyr	Asn Asn	Ile Asp	Leu Ser 1130	Leu 111: Ile	1100 Leu 5 Glu	O Asn Lys	Leu Ser	Leu Lys 113	Ile 112 Glu
Val 1109 Ala	1090 Ile 5 Ile	Phe Phe	Leu Asn Phe	Leu Asn 1125 Gln	Val 1110 Ile	1099 Ala O Tyr	Asn Asn	Ile Asp Gln	Leu Ser 1130 Leu	Leu 111: Ile	1100 Leu 5 Glu	0 Asn	Leu Ser His	Leu Lys 113! Asp	Ile 112 Glu
Val 1105 Ala Ile	1090 Ile 5 Ile Trp	Phe Phe Phe Leu	Leu Asn Phe	Leu Asn 1125 Gln	Val 1110 Ile 5 Arg	1095 Ala O Tyr Tyr	Asn Asn Asn Gln	Ile Asp Gln 114	Leu Ser 1130 Leu	Leu 111! Ile) Met	1100 Leu Glu Glu	0 Asn Lys Tyr	Leu Ser His	Leu Lys 113! Asp	Ile 112 Glu 5
Val 1105 Ala Ile	1090 Ile Ile Trp	Phe Phe Leu Leu	Leu Asn Phe 1140 Pro	Leu Asn 1125 Gln	Val 1110 Ile 5 Arg	1095 Ala O Tyr Tyr	Asn Asn Gln Ser	Ile Asp Gln 1145 Ile	Leu Ser 1130 Leu	Leu 111! Ile) Met	1100 Leu Glu Glu	O Asn Lys Tyr Val	Leu Ser His 1150 Tyr	Leu Lys 113! Asp	Ile 112 Glu 5
Val 1109 Ala Ile Pro	1090 Ile Ile Trp	Phe Phe Leu Leu 115	Leu Asn Phe 1140 Pro	Leu Asn 1125 Gln) Pro	Val 1110 Ile Arg Pro	1095 Ala Tyr Tyr	Asn Asn Gln Ser	Ile Asp Gln 1145 Ile	Leu Ser 1130 Leu 5 Phe	Leu 111! Ile) Met Ala	1100 Leu Glu Glu His	D Asn Lys Tyr Val 116	Leu Ser His 1150 Tyr	Leu Lys 113! Asp His	Ile 112 Glu Ser Phe
Val 1109 Ala Ile Pro	1090 Ile Ile Trp	Phe Phe Leu Leu 115	Leu Asn Phe 1140 Pro	Leu Asn 1125 Gln) Pro	Val 1110 Ile Arg Pro	1099 Ala Tyr Tyr Phe	Asn Asn Gln Ser 1160 Arg	Ile Asp Gln 1145 Ile	Leu Ser 1130 Leu 5 Phe	Leu 111! Ile) Met Ala	1100 Leu Glu Glu His	D Asn Lys Tyr Val 1165	Leu Ser His 1150 Tyr	Leu Lys 113! Asp His	Ile 112 Glu Ser Phe
Val 1105 Ala Ile Pro	1090 Ile Ile Trp Phe Asp	Phe Phe Leu Leu 115:	Asn Phe 1140 Pro 5	Leu Asn 1125 Gln Pro Tyr	Val 1110 Ile Arg Pro	1099 Ala Tyr Tyr Phe Leu 1179	Asn Asn Gln Ser 1160 Arg	Asp Gln 114: Ile Arg	Ser 1130 Leu 5 Phe Pro	Leu 111! Ile) Met Ala Asp	1100 Leu Glu Glu His Thr	O Asn Lys Tyr Val 1165 Lys	Leu Ser His 1150 Tyr Arg	Leu Lys 113! Asp His	Ile 112 Glu Ser Phe
Val 1105 Ala Ile Pro	1090 Ile Ile Trp Phe Asp	Phe Phe Leu Leu 115:	Leu Asn Phe 1140 Pro 5	Leu Asn 1125 Gln Pro Tyr	Val 1110 Ile Arg Pro	1099 Ala Tyr Tyr Phe Leu 1179	Asn Asn Gln Ser 1160 Arg	Asp Gln 114: Ile Arg	Ser 1130 Leu 5 Phe Pro	Leu 1119 Ile Met Ala Asp Glu	1100 Leu Glu Glu His Thr 1180 Asp	D Asn Lys Tyr Val 1165	Leu Ser His 1150 Tyr Arg	Leu Lys 113! Asp His	Ile 112 Glu Ser Phe Arg
Val 1109 Ala Ile Pro Ile Ser 1189	1090 Ile Ile Trp Phe Asp 1170 Glu	Phe Phe Leu Leu 1155 Tyr His	Asn Phe 1140 Pro Leu Ser	Asn 112: Gln Pro Tyr	Val 1110 Ile 5 Arg Pro Asn Lys 1190	Tyr Tyr Phe Leu 1179 Leu	Asn Asn Gln Ser 1160 Arg	Asp Gln 1149 Ile Arg	Ser 1130 Leu 5 Phe Pro	Leu 1119 Ile Met Ala Asp Glu 1199	Glu Glu His Thr 1180	Asn Lys Tyr Val 1169 Lys O	Leu Ser His 1150 Tyr Arg	Leu Lys 113! Asp His Phe	Ile 112 Glu Ser Phe Arg Arg 120
Val 1109 Ala Ile Pro Ile Ser 1189	1090 Ile Ile Trp Phe Asp 1170 Glu	Phe Phe Leu Leu 1155 Tyr His	Asn Phe 1140 Pro Leu Ser	Asn 112: Gln Pro Tyr	Val 1110 Ile 5 Arg Pro Asn Lys 1190	Tyr Tyr Phe Leu 1179 Leu	Asn Asn Gln Ser 1160 Arg	Asp Gln 1149 Ile Arg	Ser 1130 Leu 5 Phe Pro	Leu 1119 Ile Met Ala Asp Glu 119	Glu Glu His Thr 1180	O Asn Lys Tyr Val 1165 Lys	Leu Ser His 1150 Tyr Arg	Leu Lys 113! Asp His Phe Lys	Ile 112 Glu Ser Phe Arg 120 Arg
Val 1109 Ala Ile Pro Ile Ser 1189 Ile	Ile Ile Trp Phe Asp 1170 Glu Gln	Phe Phe Leu Leu 115: Tyr His	Leu Asn Phe 1140 Pro Leu Ser	Asn 1125 Gln Pro Tyr Ile Glu 1205	Val 1110 Ile Arg Pro Asn Lys 1190 Glu	Tyr Tyr Phe Leu 117! Leu Asp	Asn Asn Gln Ser 1160 Arg Ser Cys	Asp Gln 1145 Ile Arg Val	Ser 1130 Leu 5 Phe Pro Thr Asp 1210	Leu 1119 Met Ala Asp Glu 1199 Thr	Glu Glu His Thr 1186 Asp	Asn Lys Tyr Val 1169 Lys O Glu	Leu Ser His 1150 Tyr Arg Met	Leu Lys 113! Asp His Phe Lys Ile 121!	Ile 112 Glu Ser Phe Arg 120 Arg
Val 1109 Ala Ile Pro Ile Ser 1189 Ile	Ile Ile Trp Phe Asp 1170 Glu Gln	Phe Phe Leu Leu 115: Tyr His	Leu Asn Phe 1140 Pro Leu Ser	Asn 1125 Gln Pro Tyr Ile Glu 1205	Val 1110 Ile Arg Pro Asn Lys 1190 Glu	Tyr Tyr Phe Leu 117! Leu Asp	Asn Asn Gln Ser 1160 Arg Ser Cys	Asp Gln 1145 Ile Arg Val	Ser 1130 Leu 5 Phe Pro Thr Asp 1210	Leu 1119 Met Ala Asp Glu 1199 Thr	Glu Glu His Thr 1186 Asp	Asn Lys Tyr Val 1169 Lys O Glu	Leu Ser His 1150 Tyr Arg Met	Leu Lys 113! Asp His Phe Lys Ile 121!	Ile 112 Glu Ser Phe Arg 120 Arg
Val 1109 Ala Ile Pro Ile Ser 1189 Ile Lys	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu	Phe Phe Leu 115: Tyr His Asp	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5	Tyr Tyr Phe Leu 1179 Asp	Asn Asn Ser 1160 Arg Ser Cys	Asp Gln 1145 Ile Arg Val Ile Pro 1225	Leu Ser 1130 Leu Fhe Pro Thr Asp 1210 Leu	Leu 1119 Met Ala Asp Glu 1199 Thr	Thr 1180 Asp Leu Val	Asn Lys Tyr Val 1165 Lys Glu Thr	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230	Leu Lys 113! Asp His Phe Lys Ile 121! Leu	Ile 112 Glu 5 Ser Phe Arg 120 Arg 5 Thr
Val 1109 Ala Ile Pro Ile Ser 1189 Ile Lys	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu	Phe Phe Leu 115: Tyr His Asp	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5	Tyr Tyr Phe Leu 1179 Asp	Asn Asn Gln Ser 1160 Arg Ser Cys	Asp Gln 1149 Ile Arg Val Ile Pro 1229	Leu Ser 1130 Leu Fhe Pro Thr Asp 1210 Leu	Leu 1119 Met Ala Asp Glu 1199 Thr	Thr 1180 Asp Leu Val	Asn Lys Tyr Val 1165 Lys Glu Thr	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230	Leu Lys 113! Asp His Phe Lys Ile 121! Leu	Ile 112 Glu 5 Ser Phe Arg 120 Arg 5 Thr
Val 1109 Ala Ile Pro Ile Ser 1189 Ile Lys	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu	Phe Phe Leu Leu 1155 Tyr His Asp Lys	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5	Tyr Tyr Phe Leu 1179 Asp	Asn Asn Ser 1160 Arg Ser Cys Glu His	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp	Leu Ser 1130 Leu Fhe Pro Thr Asp 1210 Leu	Leu 1119 Met Ala Asp Glu 1199 Thr	Thr 1180 Asp Leu Val	Asn Lys Tyr Val 1169 Lys O Glu	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn	Leu Lys 113! Asp His Phe Lys Ile 121! Leu	Ile 112 Glu 5 Ser Phe Arg 120 Arg 5 Thr
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys	In the second se	Phe Leu Leu 1155 Tyr His Asp Lys Thr 1235	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn	Val 1110 Ile Arg Pro Asn Lys 1190 Glu Thr	Tyr Tyr Phe Leu 1179 Leu Asp Lys	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp	Ser 1130 Leu 5 Phe Pro Thr Asp 1210 Leu 5	Leu 1111 Ile Met Ala Asp Glu 1199 Thr	Glu His Thr 1180 Asp Leu Val	Asn Lys Tyr Val 1169 Lys Glu Thr Glu 1249	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn	Leu Lys 113! Asp His Phe Lys Ile 121! Leu Phe	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Clu Leu Leu	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn	Val 1110 Ile Arg Pro Asn Lys 1190 Glu Thr	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp	Ser 1130 Leu 5 Phe Pro Thr Asp 1210 Leu 5	Leu 1111 Ile Met Ala Asp Glu 1199 Thr	Glu Glu His Thr 1180 Asp Leu Val Gln	Asn Lys Tyr Val 1169 Lys Glu Thr Glu 1249 Asp	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn	Leu Lys 113! Asp His Phe Lys Ile 121! Leu Phe	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr
Val 1109 Ala Ile Pro Ile Ser 1189 Ile Lys Glu Leu	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu Leu Lys 1250	Phe Leu Leu 115: Tyr His Asp Lys Thr 123: Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val	Val 1110 Ile Arg Pro Asn Lys 1190 Glu Thr Arg	1099 Ala Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259	Asn Asn Ser 1160 Arg Ser Cys Glu His 1240 Ile	Asp Gln 1149 Ile Arg Val Ile Pro 1229 Asp Glu	Ser 1130 Leu Phe Pro Thr Asp 1210 Leu Thr	Leu 111! Ile Met Ala Asp Glu 119! Thr Ser Met	Glu Glu His Thr 1180 Asp Leu Val Gln Ile	Asn Lys Tyr Val 1165 Lys Glu Thr Thr Glu 1245 Asp	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His	Leu Lys 1139 Asp His Phe Lys Ile 1219 Leu Phe	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn	In the second se	Phe Leu Leu 115: Tyr His Asp Lys Thr 123: Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5 Thr Arg Tyr	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val	Asn Asn Ser 1160 Arg Ser Cys Glu His 1240 Ile	Asp Gln 1149 Ile Arg Val Ile Pro 1229 Asp Glu	Ser 1130 Leu Phe Pro Thr Asp 1210 Leu Thr	Leu 111! Ile Met Ala Asp Glu 119! Thr Ser Met Lys	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn	Asn Lys Tyr Val 1169 Lys Glu Thr Glu 1249 Asp	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His	Leu Lys 1139 Asp His Phe Lys Ile 1219 Leu Phe	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu Lys 1250 Ser	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235 Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg	Asn 1129 Gln Pro Tyr Ile Glu 1209 Asn Gln Val	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5 Thr Arg Tyr Val 1270	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val	Asn Asn Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln	Asp Gln 1145 Ile Arg Val Ile Pro 1225 Asp Glu Ile	Leu Ser 1130 Leu Phe Pro Thr Asp 1210 Leu Leu Thr	Leu 1111 Ile Met Ala Asp Glu 119 Thr Ser Met Lys Lys 127	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn	Asn Lys Tyr Val 1169 Lys Glu Thr Thr Glu 1249 Asp O Lys	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His	Leu Lys 113! Asp His Phe Lys Ile 121! Leu Phe Ile Leu	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser Ser
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260	1090 Ile Ile Trp Phe Asp 1170 Glu Glu Leu Lys 1250 Ser	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235 Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val Glu Ala	Val 1110 Ile 5 Arg Pro Asn Lys 1190 Glu 5 Thr Arg Tyr Val 1270 Ser	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val	Asn Asn Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln	Asp Gln 1145 Ile Arg Val Ile Pro 1225 Asp Glu Ile	Leu Ser 1130 Leu Phe Pro Thr Asp 1210 Leu Thr Leu Leu Leu	Leu 1111 Ile Met Ala Asp Glu 1199 Thr Ser Met Lys 127 Pro	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn	Asn Lys Tyr Val 1165 Lys Glu Thr Thr Glu 1245 Asp	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His	Leu Lys 113! Asp His Phe Lys Ile 121! Leu Phe Ile Leu Ile	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser 128 Glu
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260 Gln	In the second se	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235 Ser Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg Asp	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val Glu Ala 1285	Val 1110 11e Arg Pro Asn Lys 1190 Glu Thr Arg Tyr Val 1276 Ser	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val Ser	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln Leu	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp Glu Ile Ser	Ser 1130 Leu 5 Phe Pro Thr Asp 1210 Leu Thr Leu Leu	Leu 1111 11e 1 Met Ala Asp Glu 1191 Thr Ser Met Lys 127 Pro	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn Asp	Asn Lys Tyr Val 1169 Lys Glu Thr Glu 1249 Asp Lys Thr	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His Lys Ser	Leu Lys 113! Asp His Phe Lys 121! Leu Phe Ile Leu Ile 129	Ile 112 Glu Ser Phe Arg Arg 120 Arg 5 Thr Leu Ser 128 Glu
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260 Gln	In the second se	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235 Ser Ser	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg Asp Ala Ile	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Val Glu Ala 1285 Thr	Val 1110 11e Arg Pro Asn Lys 1190 Glu Thr Arg Tyr Val 1276 Ser	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val Ser	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln Leu	Asp Gln 1145 Ile Arg Val Ile Pro 1225 Asp Glu Ile Ser Ile	Leu Ser 1130 Leu Phe Pro Thr Asp 1210 Leu Thr Leu Leu 1290 Asp	Leu 1111 11e 1 Met Ala Asp Glu 1191 Thr Ser Met Lys 127 Pro	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn Asp	Asn Lys Tyr Val 1169 Lys Glu Thr Thr Glu 1249 Asp O Lys	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His Lys Ser	Leu Lys 113! Asp His Phe Lys Ile 121! Leu Phe Ile Leu Ile 129 Pro	Ile 112 Glu Ser Phe Arg Arg 120 Arg 5 Thr Leu Ser 128 Glu
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260 Gln Val	In the second of	Phe Phe Leu 1159 Tyr His Asp Lys Thr 1239 Ser Phe Lys	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg Asp Ala Ile 1300	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val Glu Ala 1285 Thr	Val 1110 Arg Pro Asn Lys 1190 Glu Thr Arg Tyr Val 1276 Ser Lys	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val Ser Thr	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln Leu	Asp Gln 1149 Ile Arg Val Ile Pro 1229 Asp Glu Ile Ser Ile 1309	Leu Ser 1130 Leu Pro Thr Asp 1210 Leu Thr Leu Leu 1290 Asp	Leu 1111 11e Met Ala Asp Glu 119 Thr Ser Met Lys 127 Pro 0 Cys	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn Asp His	Asn Lys Tyr Val 1169 Lys O Glu Thr Thr Glu 1249 Asp O Lys Thr	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His Lys Ser Ser	Leu Lys 113! Asp His Phe Lys 11eu Phe Leu Ile 129 Pro	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser 128 Glu 5 Val
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260 Gln Val	In the second of	Phe Phe Leu 1155 Tyr His Asp Lys Thr 1235 Ser Phe Lys Glu	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg Asp Ala Ile 1300 Asp	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val Glu Ala 1285 Thr	Val 1110 Arg Pro Asn Lys 1190 Glu Thr Arg Tyr Val 1276 Ser Lys	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val Ser Thr	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln Leu Thr	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp Glu Ile Ser Ile 130: Arg	Leu Ser 1130 Leu Pro Thr Asp 1210 Leu Thr Leu Leu 1290 Asp	Leu 1111 11e Met Ala Asp Glu 119 Thr Ser Met Lys 127 Pro 0 Cys	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn Asp His	Asn Lys Tyr Val 1169 Lys O Glu Thr Thr Glu 1249 Asp O Lys Thr Leu Leu	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His Lys Ser Ser 1310	Leu Lys 113! Asp His Phe Lys 11eu Phe Leu Ile 129 Pro	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser 128 Glu 5 Val
Val 1100 Ala Ile Pro Ile Ser 1180 Ile Lys Glu Leu Asn 1260 Gln Val	In the second of	Phe Phe Leu 1159 Tyr His Asp Lys Thr 1239 Ser Phe Lys	Leu Asn Phe 1140 Pro Leu Ser Phe Leu 1220 Cys Arg Asp Ala Ile 1300 Asp	Asn 1125 Gln Pro Tyr Ile Glu 1205 Asn Gln Val Glu Ala 1285 Thr	Val 1110 Arg Pro Asn Lys 1190 Glu Thr Arg Tyr Val 1276 Ser Lys	Tyr Tyr Phe Leu 1179 Leu Asp Lys Val Asp 1259 Val Ser Thr	Asn Asn Gln Ser 1160 Arg Ser Cys Glu His 1240 Ile Gln Leu	Asp Gln 114: Ile Arg Val Ile Pro 122: Asp Glu Ile Ser Ile 130: Arg	Leu Ser 1130 Leu Pro Thr Asp 1210 Leu Thr Leu Leu 1290 Asp	Leu 1111 11e Met Ala Asp Glu 119 Thr Ser Met Lys 127 Pro 0 Cys	Glu Glu His Thr 1180 Asp Leu Val Gln Ile 1260 Asn Asp His	Asn Lys Tyr Val 1169 Lys O Glu Thr Thr Glu 1249 Asp O Lys Thr	Leu Ser His 1150 Tyr Arg Met Arg Asp 1230 Asn His Lys Ser Ser 1310	Leu Lys 113! Asp His Phe Lys 11eu Phe Leu Ile 129 Pro	Ile 112 Glu Ser Phe Arg 120 Arg 5 Thr Leu Ser 128 Glu 5 Val

PCT/US99/29996 WO 00/40614

-25-

Gln Arg Asp His Thr Leu Arg Lys Leu Pro Thr Trp Glu Thr Ser Thr Ala Ser Thr Ser Ser Phe Glu Phe Val Phe Tyr Phe Thr Arg His Glu 1350 1355 Gly Asn Glu Asn Lys Tyr Glu Phe Lys Lys Leu Glu Lys Gly Gly Phe Trp Arg Asn Asn Tyr Val Ile Ser Trp Arg Leu

<210> 15 <211> 1868 <212> PRT <213> C. Elegans

<400> 15 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln Ser Lys Gly Gly Asp Gln Val Pro Pro Thr Ser Thr Thr Thr Gly Gly Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val 85 90 95 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu . 120 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr 275 280 285 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Leu Lys Lys Arg Glu Asp Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Lys Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

						-		•	-26-						
		355					360					365			
Asp	Asn 370	Gly	Thr	Val	Gly	Arg 375	Tyr	Gly	Ala	Glu	Val 380	Ile	Leu	Arg	Lys
Arg 385		Glu	Met	Tyr	11e 390	Ser	Gln	Lys	Gln	Lys 395	Ile	Phe	Gly	Gly	Thr 400
Arg	Ser	Val	Pro	Val 405	Val	Cys	Val	Val	Leu 410	Glu	Gly	Gly	Ser	Cys 415	Thr
Ile	Arg	Ser	Val 420	Leu	Asp	Tyr	Val	Thr 425	Asn	Val	Pro	Arg	Val 430	Pro	Val
Val	Val	Cys 435	Asp	Gly	Ser	Gly	Arg 440	Ala	Ala	Asp	Leu	Leu 445	Ala	Phe	Ala
His	Gln 450		Val	Thr	Glu	Asp 455	Gly	Leu	Leu	Pro	Asp 460	Asp	Ile	Arg	Arg
Gln 465	Val	Leu	Leu	Leu	Val 470	Glu	Thr	Thr	Phe	Gly 475	Cys	Ser	Glu	Ala	Ala 480
				485					490					Lys 495	
			500					505					510	Val	
		515					520					525		Ala	
-	530					535					540			Ala	
545	_				550					555				Leu	560
				565					570					Val 575	
			580					585					590	Ile	
		595					600					605		Thr	
	610					615					620			Arg	
625					630					635				Asn	640
				645					650					Lys 655	
			660					665					670	Val	
		675					680					685		Gln Gly	
	690					695					700			His	
705					710					715				Ile	720
				725					730					735 Arg	
_	_		740					745					750		
		755					760					765		Glu	
	770					775					780			Asp	
785	_				790					795				Lys	800
				805					810					Met 815	
			820					825					830	Glu	
Ата	GLU	Asp	ıyr	ьеи	GIU	vaı	ьти	тте	Cys	GTU	GIU	neu	rN_{2}	Lys	TÀT

								•	-21-						
		835					840					845			
Ala	Glu 850	Glu	Phe	Arg	Ile	Leu 855	Ser	Leu	Glu	Leu	Leu 860	Asp	His	Cys	Tyr
His \ 865	Val	Asp	Asp	Ala	Gln 870	Thr	Leu	Gln	Leu	Leu 875	Thr	Tyr	Glu	Leu	Ser 880
Asn :	Trp	Ser	Asn	Glu 885	Thr	Cys	Leu	Ala	Leu 890	Ala	Val	Ile	Val	Asn 895	Asn
Lys 1	His	Phe	Leu 900	Ala	His	Pro	Cys	Cys 905	Gln	Ile	Leu	Leu	Ala 910	Asp	Leu
Trp 1	His	Gly 915	Gly	Leu	Arg	Met	Arg 920	Thr	His	Ser	Asn	Ile 925	Lys	Val	Val
Leu (Gly 930	Leu	Ile	Cys	Pro	Pro 935	Phe	Ile	Gln	Met	Leu 940	Glu	Phe	Lys	Thr
Arg (945					950					955					960
Asp i	Met	Asn	Tyr	Ser 965	Ser	Ser	Ser	Ser	Ser 970	Ser	Ser	Ser	Ser	Ser 975	Ser
Ser S	Ser	Ser	Ser 980	Ser	Asp	Ser	Ser	Ser 985	Phe	Glu	Asp	Asp	Asp 990	Asp	Glu
Asn A	Asn	Ala 995	His	Asn	His	Asp	Gln 1000		Arg	Thr	Arg	Lys 1005		Ser	Gln
Gly S	Ser 1010		Gln	Ser	Leu	Asn 1015		Thr	Ser	Leu	Phe 1020		Ser	Arg	Arg
Arg 1 1025	Lys	Ala	Lys	Lys	Asn 1030		Lys	Cys	Asp	Arg 1035		Thr	Asp	Ala	Ser 104
Ala(_			1045	5				1050)				1055	5
Glu '	-		1060)				1065	5				1070)	
Tyr I		1075	5				1080)				1085	5		
	1090)				1095	5				1100)			
Leu (1110)				1115	5				112
Gln 1				1125	5				1130)				1135	5
Phe '			1140)				1145	5				1150)	
Ile 1		1155	5				1160)				1165	5		
-	1170)				1175	5				1180)			
Phe (_				1190) _	_			1199	5				120
Phe :				1205	5				1210)				1215	5
Thr I			1220)				1225	5				1230)	
Phe (1235	5				1240)				1245	5		
	1250	1				1255	5				1260)			
Gly 1 1265					1270)				1275	5				128
Ile :				1285)				1290)				1295	5
Gln S			1300)				1305	5				1310)	
Asn :	Ile	Phe	Leu	Lys	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala

PCT/US99/29996 WO 00/40614 -28-

					-	28 ~						
1315	5			1320)				1325			
Asp Glu Ile 1330			1335)				1340)			
Asn Gly Gly	Pro Val	Ile 1350	Leu	Gly	Asn	Gly	Thr 1355	Thr	Gly	Leu	Ser	Cys 136
Val Pro Gly	Tyr Trp	Ile		Pro	Leu	Leu 1370	Met		Phe	Phe	Leu 1375	
Ile Ala Asn			Met	Ser	Met 1385	Leu		Ala	Ile	Phe 1390	Asn	
Ile Phe Asp	Ala Thr	Asp		Met 1400	Ser		Gln	Ile	Trp 1405	Leu		Gln
1395 Arg Tyr Lys 1410	Gln Val			Tyr		Ser	Thr	Pro 1420	Phe		Pro	Pro
Pro Leu Thr	Pro Leu		His		Val	Leu	Ile 1435	Leu		Phe	Val	Arg 144
Thr Arg Leu	Ser Cys	Ser		Ser	Gln	Glu 1450	Arg		Pro	Ile	Leu 1455	Leu
Leu Lys Ile			Phe		Asp 1465	Asn		Gln	Ile	Glu 1470	Lys	
His Asp Phe	Glu Glu	Asp			Glu		Leu	Ala	Arg 1485	Gln		Leu
Asn Glu Lys	Asn Thr			Glu		Arg	Ile	Leu 1500	Arg		Asp	Ile
Arg Thr Asp	Gln Ile		Asn		Leu	Ile	Asp 1515	Leu		Ala	Lys	Glu 152
Ser Met Gly	Arg Asp	Val	Ile	Asn	Asp	Val 1530	Glu		Arg	Leu	Ala 1535	Ser
Val Glu Lys	Ala Gln 1540	Asn	Glu		Leu 1545	Glu		Val	Arg	Ala 1550		Leu
Asn Gln Asn 1555	Asn Ala	Pro	Thr		Ile		Arg	Cys	Phe 1565		Pro	Ser
Pro Asp Pro 1570			Thr 1575	Ala		Gly	Thr	Pro 1580		Pro	Leu	Leu
Leu Lys Leu 1585	Pro Gly		Asp		Ile	Leu	Glu 1595	Glu	Lys	Asp	His	Asp 160
Ser Gly Glu	1609	5				1610)				1615	5
Arg Thr Ala	1620				1625	5				1630)	
Met Leu Leu 1635	5			1640)				1645	5		
Arg Leu Met 1650			1655	,				1660)			
Lys Leu Ser 1665		1670					1675	,				168
Ile Thr Asp	1689	5				1690)				1695	5
Asn Arg Ser	1700				1705	5				1710)	
Gly Gly Gly 1715	.			1720)				1725	5		
Leu Arg Met 1730	Thr Pro		Ser 1735		Val	Glu	Glu	Ser 1740	Thr	Ser	Arg	Asp
Gln Ile Phe 1745		1750					17:55	j				176
Ala Asp Cys	Glu Leu 1765		Asp	Val	Ile	Thr 1770		Glu	Glu	Asp	Glu 1775	Glu
Glu Asp Asp	Glu Glu 1780	Asp .			1785	5				1790)	
Arg Arg Lys		Arg	Gln	Asn	Arg	Gln	Pro	Ser	His	Thr	Leu	Glu

-29-

```
1800
                                             1805
Thr Asp Leu Ser Glu Gly Glu Glu Val Asp Pro Leu Asp Val Leu Lys
   1810 1815
                                         1820
Met Lys Glu Leu Pro Ile Ile His Gln Ile Leu Asn Glu Glu Gln
                                    1835
1825 1830
Ala Gly Ala Pro His Ser Thr Pro Val Ile Ala Ser Pro Ser Ser Ser
                                1850
           1845
Arg Ala Asp Leu Thr Ser Gln Lys Cys Ser Asp Val
           1860
                     1865
     <210> 16
     <211> 489
     <212> DNA
     <213> Mus Musculus
     <400> 16
ccctgaaaga ctcgacttct gctgctagcg ctggagctga gttagttttg agaaggtttc
                                                                    120
ceqqqqctqt cettqttcqq tqqcccqtqc caccqcctcc ggaqacqctt tccgatagat
ggctgcaggc cgcggaggtg gaggaggagc cgctgccctt ccggagtccg ccccgtgagg
                                                                    180
agaatgtccc agaaatcctg gatagagagc actttgacca agagggagtg tgtatatatt ataccaagct ccaaagacc tcacagatgt cttccaggat gtcagatttg tcagcaactt
                                                                    240
                                                                    300
gteagatgtt tetgtggteg tttggteaag caacatgcat getttaetge aagtettgee
                                                                    360
atgaaatact cagatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct
                                                                    420
gtggaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg
                                                                    480
                                                                    489
ggttctcat
     <210> 17
     <211> 102
     <212> PRT
     <213> Mus Musculus
     <400> 17
Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys
                                10
Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly
         20
                               25
                                                  30
Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val
      35
                         40
Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp
                     55
Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val
                 70
                                    75
Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn
                                  90
              85
Phe Gln Gly Gly Ser His
           100
     <210> 18
     <211> 410
     <212> DNA
     <213> Homo Sapiens
     <220>
     <221> unsure
     <222> (6)...(6)
     <221> unsure
     <222> (58)...(58)
     <221> unsure
```

60

60

-30-

```
<222> (89)...(89)
       <221> unsure
       <222> (406) ... (406)
      <400> 18
 gccgcnggag cctgagcgga gggtgtgcgc agcctcgcca gcgggggccc cgggctgngc
 cattgcctca ctgagccagc gcctgcctnc tacctcgccg acagctggaa ccagtgcgac
                                                                       120
 ctaqtqqctc tcacctgctt cctcctgggc gtgggctgcc ggctgacccc gggtttgtac
                                                                       180
 cacctgggcc gcactgtcct ctgcatcgac ttcatggttt tcacggtgcg gctgcttcac
                                                                       240
                                                                       300
 atcttcacqq tcaacaaaca gctggggccc aagatcgtca tcgtgagcaa gatgatgaag
                                                                       360
 gacgtgttct tetteetett etteetegge gtgtggetgg tagetatggg ttgggeeaeg
gaggggttcc tgaggccacg ggacagtgac ttcccaagta tcctgncgcc
                                                                       410
      <210> 19
       <211> 131
       <212> PRT
       <213> Homo Sapiens
      <220>
      <221> UNSURE
       <222> (15)...(15)
      <223> UNKNOWN
      <221> UNSURE
      <222> (25)...(25)
      <223> UNKNOWN
      <221> UNSURE
       <222> (131)...(131)
      <223> UNKNOWN
      <400> 19
Ala Glu Gly Val Arg Ser Leu Ala Ser Gly Gly Pro Gly Leu Xaa His
                                   10
. 1
                 5
Cys Leu Thr Glu Pro Ala Pro Ala Xaa Tyr Leu Ala Asp Ser Trp Asn
                                                     30
                                25
            20
Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys
                                               45
                           40
        35
Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile
                        55
                                             60
Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn
                                        75
                    70
Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp
                                    90
                85
Val Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Met Gly
            100
                                105
                                                     110
Trp Ala Thr Glu Gly Phe Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser
        115
                            120
Ile Leu Xaa
    130
      <210> 20
      <211> 389
      <212> DNA
      <213> Homo Sapiens
      <400> 20
caaatttttt gttagtacac catctcatcc aaattgcaaa agtcacatgg aaactggaac
                                                                       120
caaagatcaa gaaactgttt gctctaaagc tacagaagga gataatacag aatttggagc
```

-31-

```
atttgtagga cacagagata gcatggattt acagaggttt aaagaaacat caaacaagat
                                                                    180
aaaaatacta tocaataaca atacttotga aaacactttg aaacgagtga gttotottgo
                                                                    240
tggatttact gactgtcaca gaacttccat tcctgttcat tcaaaacgag aaaagatcag
                                                                    300
tagaaqqcca tctaccqaaq acactcatga aqtagattcc aaaqcaqctt taataccggt
                                                                    360
                                                                    389
ttgtagattt caactaaaca gatatatat
      <210> 21
      <211> 415
      <212> DNA
      <213> Homo Sapiens
      <400> 21
atttctagtt tttcaaattt gccagtcttt ttgaatagta tctccttctt ttctcatgtt
                                                                     60
ttatatttaa aactttttaa tgtccatcat cactttaaac atacttattt tgtcatctat
                                                                    120
aaccaataat tocactatot tatoagaaat caaatacogt ttatgtaagt tgactoocat
                                                                    180
gagttctaaa ttgccattgt gaggtcatct tcggttaggc tttaatttgt tgcaaagttg
                                                                    240
tgcaqctcaq ggtcaggaag agtccctcca gaaaggagga tttgttactg tgaatctctt
                                                                    300
                                                                    360
tgttaactaa cctctttccc cactgaaata acttttttca ataacatgat tttaacaaca
taatctctct atgccagaac agatatatat gaatgtaagt caatattttc ttgag
                                                                    415
      <210> 22
      <211> 405
      <212> DNA
      <213> Mus Musculus
      <400> 22
ttattatggc ttatcatgaa aaaccagtcc tgcctcctcc tcttatcatc ctcagccata
                                                                     60
                                                                    120
tagtttcact gttttgctgt gtatgcaaaa gaagaaagaa agataagact tccgatgggc
caaaactttt cttaacagaa gaagatcaaa agaaactcca tgattttgaa gagcagtgtg
                                                                    180
                                                                    240
gggtcacttt tgaaagagtg gagcagatga gcattcagat taaagaagtt ggagatcgtg
                                                                    300
tcaactacat aaaaagatca ttacagtctt tagattctca aattggtcat ctgcaagatc
                                                                    360
tctcaqccct aacagtagat acattgaaaa cacttacagc ccaga
                                                                    405
      <210> 23
      <211> 5117
      <212> DNA
      <213> Homo Sapiens
     <220>
     <221> unsure
     <222> (2382)...(2382)
     <223> unknown
     <221> unsure
     <222> (4664)...(4664)
     <223> unknown
     <221> unsure
     <222> (4682)...(4682)
     <223> unknown
     <221> unsure
     <222> (4702)...(4702)
     <223> unknown
     <221> unsure
     <222> (5038)...(5039)
     <223> unknown
```

<221> unsure <222> (5056)...(5056) <223> unknown <221> unsure <222> (5071)...(5072)

<400> 23

60 gatggcaaca tggtgaagaa tcaatggcta aagcattagt tgcctgtaag atctatcgtt caatggcata tgaagcaaag cagagtgacc tggtagatga tacttcagaa gaactaaaac 120 agtattccaa tgattttggt cagttggccg ttgaattatt agaacagtcc ttcagacaag 180 240 atgaaaccat ggctatgaaa ttgctcactt atgaactgaa gaactggagt aattcaacct qccttaaqtt agcagtttct tcaagactta gaccttttgt agctcacacc tgtacacaaa 300 tgttgttatc tgatatgtgg atgggaaggc tgaatatgag gaaaaattcc tggtacaagg 360 tcatactaag cattttagtt ccacctgcca tattgctgtt agagtataaa actaaggctg 420 480 aaatgtccca tatcccacaa tctcaagatg ctcatcagat gacaatggat gacagcgaaa acaactttca gaacataaca gaagagatcc ccatggaagt gtttaaagaa gtacggattt 540 600 tggatagtaa tgaaggaaag aatgagatgg agatacaaat gaaatcaaaa aagcttccaa ttacgcgaaa gttttatgcc ttttatcatg caccaattgt aaaattctgg tttaacacgt 660 tggcatattt aggatttctg atgctttata catttgtggt tcttgtacaa atggaacagt 720 taccttcagt tcaagaatgg attgttattg cttatatttt tacttatgcc attgagaaag 780 840 tccqtqaqat ctttatgtct gaagctggga aagtaaacca gaagattaaa gtatggttta gtgattactt caacatcagt gatacaattg ccataatttc tttcttcatt ggatttggac 900 taagatttgg agcaaaatgg aactttgcaa atgcatatga taatcatgtt tttgtggctg 960 gaagattaat ttactgtctt aacataatat tttggtatgt gcgtttgcta gattttctag 1020 1080 ctgtaaatca acaggcagga ccttatgtaa tgatgattgg aaaaatggtg gccaatatgt tctacattgt agtgattatg gctcttgtat tacttagttt tggtgttccc agaaaggcaa 1140 1200 tactttatcc tcatgaagca ccatcttgga ctcttgctaa agatatagtt tttcacccat actggatgat ttttggtgaa gtttatgcat acgaaattga tgtgtgtgca aatgattctg 1260 1320 ttatccctca aatctqtqqt cctqqqacqt ggttgactcc atttcttcaa gcagtctacc totttgtaca gtatatoatt atggttaato ttottattgc atttttcaac aatgtgtatt 1380 1440 tacaagtgaa ggcaatttcc aatattgtat ggaagtacca gcgttatcat tttattatgg cttatcatga gaaaccagtt ctgcctcctc cacttatcat tcttagccat atagtttctc 1500 1560 tgttttgctg catatgtaag agaagaaaga aagataagac ttccgatgga ccaaaacttt tottaacaga agaagatcaa aagaaactto atgattttga agagcagtgt gttgaaatgt 1620 atttcaatga aaaagatgac aaatttcatt ctgggagtga agagagaatt cgtgtcactt 1680 ttgaaagagt ggaacagatg tgcattcaga ttaaagaagt tggagatcgt gtcaactaca 1740 1800 taaaaagatc attacaatca ttagattctc aaattggcca tttgcaagat ctttcagccc tgacggtaga tacattaaaa acactcactg cccagaaagc gtcggaagct agcaaagttc 1860 ataatgaaat cacacgagaa ctgagcattt ccaaacactt ggctcaaaac cttattgatg 1920 atggtcctgt aagaccttct gtatggaaaa agcatggtgt tgtaaataca cttagctcct 1980 ctcttcctca aggtgatctt gaaagtaata atccttttca ttgtaatatt ttaatgaaag 2040 atgacaaaga tccccagtgt aatatatttg gtcaagactt acctgcagta ccccagagaa 2100 aagaatttaa ttttccagag gctggttcct cttctggtgc cttattccca agtgctgttt 2160 cccctccaga actgcgacag agactacatg gggtagaact cttaaaaaata tttaataaaa 2220 atcaaaaatt aggcagttca tctactagca taccacatct gtcatcccca ccaaccaaat 2280 2340 tttttgttag tacaccatct cagccaagtt gcaaaagcca cttggaaact ggaaccaaag atcaagaaac tgtttgctct aaagctacag aaggagataa tncagaattt ggagcatttg 2400 taggacacag agatagcatg gatttacaga ggtttaaaga aacatcaaac aagataaaaa 2460 tactatecaa taacaatact tetgaaaaca etttgaaacg agtgagttet ettgetggat 2520 2580 ttactgactg tcacagaact tccattcctg ttcattcaaa acaagcagaa aaaatcagta gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta ataccggatt 2640 ggttacaaga tagaccatca aacagagaaa tgccatctga agaaggaaca ttaaatggtc 2700 tcacttctcc atttaagcca gctatggata caaattacta ttattcagct gtggaaagaa 2760 2820 ataacttgat gaggttatca cagagcattc catttacacc tgtgcctcca agaggggagc ctgtcacagt gtatcgtttg gaagagagtt cacccaacat actaaataac agcatgtctt 2880 2940 cttggtcaca actaggcctc tgtgccaaaa tagagttttt aagcaaagag gagatgggag 3000 gaggtttacg aagagctgtc aaagtacagt gtacctggtc agaacatgat atcctcaaat 3060 cagggcatct ttatattatc aaatcttttc ttccagaggt ggttaataca tggtcaagta 3120 tttataaaga agatacagtt ctgcatctct gtctgagaga aattcaacaa cagagagcag

-33cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa ggttccttga agttttcctg ctgtattgcc attcagcagg acagtggttt gctgtggaag 3240 3300 aatgtatgac tggagaattt agaaaataca acaataataa tggagatgag attattccaa ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag 3360 3420 gggagttact ggtacttgat ttgcaaggtg ttggtgaaaa tttgactgac ccatctgtga taaaagcaga agaaaagaga tcctgtgata tggtttttgg cccagcaaat ctaggagaag 3480 3540 atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttcct caggatgagc 3600 cttcagattt gaatcttcag cctggaaatt ccaccaaaga atcagaatca gctaattctg 3660 ttcgtctgat gttataatat taatattact gaatcattgg ttttgcctgc acctcacaga 3720 aatgttactg tgtcactttt ccctcgggag gaaattgttt ggtaatatag aaaggtgtat 3780 gcaagttgaa tttgctgact ccagcacagt taaaaggtca atattctttt gacctgatta 3840 atcagtcaga aagtccctat aggatagagc tggcagctga gaaattttaa aggtaattga 3900 3960 taattagtat ttgtaacttt ttaaagggct ctttgtatag cagaggatct catttgactt tgttttgatg agggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt 4020 4080 gtccatgcag attttattgg cagctgtttt attgccattc aactagggaa atgaagaaat cacgcagcct tttggttaaa tggcagtcaa aattttcctc agtgtattta gtgtgttcag 4140 tgatgatatc actggttccc aactagatgc ttgttggcca cgggaaggga aatgacttgt 4200 tctaattcta ggttcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga 4260 4320 cggtatttat gaatcagggt taggctccat atttaaagat agagccagtt tttttttaa atagaaccca aattgtgtaa aaatgttaat tgggtttttt aaacattgtt ttatcaagtc 4380 4440 actgttaagt agaagaaagc catggtaaac tgatacataa cctaaattat aaaagcagaa acctaactca ctcgtcaagg gaagttacct tttgaggaaa gttaaagtac ttttttccct 4500 4560 atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt 4620 tatatcaaat cagaatgtaa acatgaactc ttgcatatat ttaaaattgt gttggaacat 4680 ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnggatta tcattatgtg anactggcag attgcagtgc ancettatgc caataaaatg taatttaaca geeceagata 4740 4800 ttgttgaata ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt ttttctttt ttttctttt cttttgtttc cttggtacta attttaattt ttatttggaa 4860 gggagcagta taaagcttat ttgtatttag tagtgtatct catagataca gacaaggcaa 4920 gagatgataa gctgtttaaa tagtgtttaa tattgattgg gggtggggag aaagaaaaag 4980 5040 tgtattactt aaagatacta tatacgtttt gtatatcatt aaatctttaa aagaaatnna ataaatttat tgtttncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag 5100 5117 gatecetega ggggeee

```
<210> 24
      <211> 1224
      <212> PRT
      <213> Homo Sapiens
      <220>
      <221> UNSURE
      <222> (794)...(794)
      <223> UNKNOWN
      <400> 24
Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys
                 5
Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp
                                25
                                                    30
           20
Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu
                                                45
                            40
       35
Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala
                        55
Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys
                                        75
                    70
Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr
```

-34-Cys Thr Gln Met Leu Leu Ser Asp Met Trp Met Gly Arg Leu Asn Met Arg Lys Asn Ser Trp Tyr Lys Val Ile Leu Ser Ile Leu Val Pro Pro Ala Ile Leu Leu Leu Glu Tyr Lys Thr Lys Ala Glu Met Ser His Ile 130 135 Pro Gln Ser Gln Asp Ala His Gln Met Thr Met Asp Asp Ser Glu Asn Asn Phe Gln Asn Ile Thr Glu Glu Ile Pro Met Glu Val Phe Lys Glu Val Arg Ile Leu Asp Ser Asn Glu Gly Lys Asn Glu Met Glu Ile Gln 180 185 190 Met Lys Ser Lys Lys Leu Pro Ile Thr Arg Lys Phe Tyr Ala Phe Tyr His Ala Pro Ile Val Lys Phe Trp Phe Asn Thr Leu Ala Tyr Leu Gly Phe Leu Met Leu Tyr Thr Phe Val Val Leu Val Gln Met Glu Gln Leu Pro Ser Val Gln Glu Trp Ile Val Ile Ala Tyr Ile Phe Thr Tyr Ala Ile Glu Lys Val Arg Glu Ile Phe Met Ser Glu Ala Gly Lys Val Asn Gln Lys Ile Lys Val Trp Phe Ser Asp Tyr Phe Asn Ile Ser Asp Thr Ile Ala Ile Ile Ser Phe Phe Ile Gly Phe Gly Leu Arg Phe Gly Ala Lys Trp Asn Phe Ala Asn Ala Tyr Asp Asn His Val Phe Val Ala Gly Arg Leu Ile Tyr Cys Leu Asn Ile Ile Phe Trp Tyr Val Arg Leu Leu Asp Phe Leu Ala Val Asn Gln Gln Ala Gly Pro Tyr Val Met Met Ile Gly Lys Met Val Ala Asn Met Phe Tyr Ile Val Val Ile Met Ala Leu Val Leu Leu Ser Phe Gly Val Pro Arg Lys Ala Ile Leu Tyr Pro His Glu Ala Pro Ser Trp Thr Leu Ala Lys Asp Ile Val Phe His Pro Tyr Trp Met Ile Phe Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys Ala Asn Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu Thr 420 425 430 Pro Phe Leu Gln Ala Val Tyr Leu Phe Val Gln Tyr Ile Ile Met Val Asn Leu Leu Ile Ala Phe Phe Asn Asn Val Tyr Leu Gln Val Lys Ala Ile Ser Asn Ile Val Trp Lys Tyr Gln Arg Tyr His Phe Ile Met Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp Lys 500 505 510 Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu Lys Asp Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr Phe Glu Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp Arg

WO 00/40614 PCT/US99/29996 -35-

									-35-						
Val	Asn	Tyr	Ile 580	Lys	Arg	Ser	Leu	Gln 585	Ser	Leu	Asp	Ser	Gln 590	Ile	Gly
		595	•				600			_		605	-	Thr	
	610					615					620			Ile	
625					630					635				Asp	640
_			_	645					650					Asn 655	
			660					665					670	Pro	
		675					680					685		Asn	
	690					695					700			Asn Val	
705					710					715				Lys	720
				725					730					735 Pro	
			740					745					750	Gln	
		755					760					765		Thr	
	770 Ser	Lys	Ala	Thr		775 Gly	Asp	Asn	Xaa	Glu	780 Phe	Gly	Ala	Phe	
785 Gly	His	Arg	Asp		790 Met	Asp	Leu	Gln		795 Phe	Lys	Glu	Thr	Ser	800 Asn
Lys	Ile	Lys	Ile 820	805 Leu	Ser	Asn	Asn	Asn 825	810 Thr	Ser	Glu	Asn	Thr 830	815 Leu	Lys
Arg	Val	Ser 835		Leu	Ala	Gly	Phe 840		Asp	Cys	His	Arg 845		Ser	Ile
	850	His				855					860			Ser	
865					870					875				Asp	880
				885					890					Gly 895	
		_	900					905					910	Asn	
		915					920					925		Gln Val	
	930					935					940			Ser	
945					950					955				Lys	960
				965					970					975 Thr	
			980					985					990	Lys	
	Leu	995 Pro				Asn	1000 Thr)			Ile	1005 Tyr	5	Glu	
			His	Leu				Glu	Ile				Arg	Ala	
1025 Gln		Leu	Thr	Phe 1045			Asn	Gln	Met 1050			Lys	Ser	Ile 1059	
				107.	-					-					-

-36-

Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala 1065 1060 Gly Gln Tro Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys 1080 1085 1075 Tyr Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu 1095 1100 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly 1110 1115 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp 1130 1135 1125 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe 1145 1140 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys 1160 1165 1155 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys 1175 1180 1170 Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro 1190 1195 Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser 1205 1210 Ala Asn Ser Val Arg Leu Met Leu 1220 <210> 25

<210> 25 <211> 2180 <212> DNA <213> Homo Sapiens

<400> 25

60 tcgaggccaa gaattcggca cgagggcctc gggcaggccc cctggagcga cctgcttctt tgggcactgt tgctgaacag ggcacagatg gccatgtact tctgggagat gggttccaat 120 gcagtttcct cagetettgg ggcctgtttg ctgctccggg tgatggcacg cctggagcct 180 240 qacqctqaqq aggcaqcacq gaggaaagac ctggcgttca agtttgaggg gatgggcgtt gacetetttg gegagtgeta tegeageagt gaggtgaggg etgeeegeet ceteeteegt 300 cyctycccyc tetgygygya tyccaettyc etccayetyy ceatycaaye tyacycccyt 360 gccttctttg cccaggatgg ggtacagtct ctgctgacac agaagtggtg gggagatatg 420 480 gecageacta cacceatetg ggecetggtt etegeettet tttgecetee acteatetae acccgcctca tcaccttcag gaaatcagaa gaggagccca cacgggagga gctagagttt 540 600 gacatggata gtgtcattaa tggggaaggg cctgtcggga cggcggaccc agccgagaag 660 acgccgctgg gggtcccgcg ccagtcgggc cgtccgggtt gctgcggggg ccgctgcggg 720 qqqcqccqqt qcctacqccq ctqqttccac ttctqqqqcq cqccqqtqac catcttcatq 780 ggcaacgtgg tcagctacct gctgttcctg ctgcttttct cgcgggtgct gctcgtggat 840 ttccagccgg cgccgcccgg ctccctggag ctgctgctct atttctgggc tttcacgctg 900 ctgtgcgagg aactgcgcca gggcctgagc ggaggcgggg gcagcctcgc cagcgggggc 960 cccgggcctg gccatgcctc actgagccag cgcctgcgcc tctacctcgc cgacagctgg 1020 aaccagtgcg acctagtggc teteacetgc tteeteetgg gegtgggetg eeggetgace ccgggtttgt accacctggg ccgcactgtc ctctgcatcg acttcatggt tttcacggtg 1080 1140 cggctgcttc acatcttcac ggtcaacaaa cagctggggc ccaagatcgt catcgtgagc 1200 aaqatqatqa aqqacqtqtt cttcttcctc ttcttcctcg gcgtgtggct ggtagcctat ggcgtggcca cggaggggct cctgaggcca cgggacagtg acttcccaag tatcctgcgc 1260 cgcgtcttct accgtcccta cctgcagatc ttcgggcaga ttccccagga ggacatggac 1320 gtggccctca tggagcacag caactgctcg tcggagcccg gcttctgggc acaccctcct 1380 ggggcccagg cgggcacctg cgtctcccag tatgccaact ggctggtggt gctgctcctc 1440 gtcatcttcc tgctcgtggc caacatcctg ctggtcaact tgctcattgc catgttcagt 1500 1560 tacacattcg gcaaagtaca gggcaacagc gatctctact ggaaggcgca gcgttaccgc 1620 ctcatccggg aattccactc tcggcccgcg ctggcccgc cctttatcgt catctcccac ttgcgcctcc tgctcaggca attgtgcagg cgacccsgga gcccccagcc gtcctccccg 1680 1740 qccctcqaqc atttccqqqt ttacctttct aaggaagccg agcggaagct gctaacgtgg 1800 gaatcggtgc ataaggagaa ctttctgctg gcacgcgcta gggacaagcg ggagagcgac tecgagmgte tgaagegeae gteecagaag gtggaettgg caetgaaaca getgggaeae 1860

1920 1980

2040

2100

2160 2180

```
-37-
atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg
gcccccgcac ctggtggcct tgtccttgag gtgagcccca tgtccatctg ggccactgtc
aggaccacct ttgggagtgt catccttaca aaccacagca tgcccggctc ctcccagaac
cagteceage etgggaggat caaggeetgg atccerggee gttatecate tggaggetge
agggteettg gggtaacagg gaccacagae eceteaceae teacagatte etcacactgg
ggaaataaag ccatttcaga
     <210> 26
     <211> 725
     <212> PRT
     <213> Homo Sapiens
     <220>
     <221> UNSURE
     <222> (553)...(553)
     <223> UNKNOWN
     <221> UNSURE
     <222> (603)...(603)
     <223> UNKNOWN
     <400> 26
Ser Arg Pro Arg Ile Arg His Glu Gly Leu Gly Gln Ala Pro Trp Ser
                                 10
Asp Leu Leu Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met
                           25
     20
Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala
                         40
     35
Cys Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu
                    55
Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
                  70
                                     75
Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg
                               90
              85
Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln
                           105
                                                 110
         100
Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val
                                           125
                        120
       115
Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr
                                        140
                     135
Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr
                 150
                                   155
Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu
                                 170
              165
Glu Leu Glu Phe Asp Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val
                            185
                                                190
           180
Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln
                        200
                                             205
      195
Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys
                     215
                                        220
   210
Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met
                                     235
                  230
Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Phe Ser Arg Val
               245
                                250
Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu
                             265
                                                270
Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly
                          280
                                             285
Leu Ser Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly
```

300

His 305	Ala	Ser	Leu	Ser	Gln 310	Arg	Leu	Arg	Leu	Tyr 315	Leu	Ala	Asp	Ser	Trp 320
Asn	Gln	Cys	Asp	Leu 325	Val	Ala	Leu	Thr	Cys 330	Phe	Leu	Leu	Gly	Val 335	Gly
Cys	Arg	Leu	Thr 340	Pro	Gly	Leu	Tyr	His 345	Leu	Gly	Arg	Thr	Val 350	Leu	Cys
lle	Asp	Phe 355	Met	Val	Phe	Thr	Val 360		Leu	Leu	His	11e 365		Thr	Val
Asn	Lys 370		Leu	Gly	Pro	Lys 375		Val	Ile	Val	Ser 380		Met	Met	Lys
Asp 385		Phe	Phe	Phe	Leu 390		Phe	Leu	Gly	Val 395		Leu	Val	Ala	Tyr 400
Gly	Val	Ala	Thr	Glu 405		Leu	Leu	Arg	Pro 410		Asp	Ser	Asp	Phe 415	
Ser	Ile	Leu	Arg 420		Val	Phe	Tyr	Arg 425		Tyr	Leu	Gln	Ile 430		Gly
Gln	Ile	Pro 435	Gln	Glu	Asp	Met	Asp		Ala	Leu	Met	Glu 445		Ser	Asn
Cys	Ser 450	Ser	Glu	Pro	Gly	Phe 455		Ala	His	Pro	Pro		Ala	Gln	Ala
Gly 465		Cys	Val	Ser	Gln 470		Ala	Asn	Trp	Leu 475		Val	Leu	Leu	Leu 480
	Ile	Phe	Leu	Leu 485		Ala	Asn	Ile	Leu 490		Val	Asn	Leu	Leu 495	
Ala	Met	Phe	Ser 500		Thr	Phe	Gly	Lys 505		Gln	Gly	Asn	Ser 510		Leu
Tyr	Trp	Lys 515	Ala	Gln	Arg	Tyr	Arg 520		Ile	Arg	Glu	Phe 525		Ser	Arg
Pro	Ala 530	Leu	Ala	Pro	Pro	Phe 535		Val	Ile	Ser	His 540		Arg	Leu	Leu
Leu 545		Gln	Leu	Cys	Arg 550		Pro	Xaa	Ser	Pro 555	Gln	Pro	Ser	Ser	Pro 560
Ala	Leu	Glu	His	Phe 565		Val	Tyr	Leu	Ser 570		Glu	Ala	Glu	Arg 575	Lys
Leu	Leu	Thr	Trp 580	Glu	Ser	Val	His	Lys 585	Glu	Asn	Phe	Leu	Leu 590	Ala	Arg
Ala	Arg	Asp 595	Lys	Arg	Glu	Ser	Asp 600	Ser	Glu	Xaa	Leu	Lys 605	Arg	Thr	Ser
	610		Asp			615					620				
625			Leu		630					635					640
			Pro	645					650					655	
			Val 660					665					670		
		675	Gly				680					685			
	690		Pro			695					700				
Val 705	Thr	Gly	Thr	Thr	Asp 710	Pro	Ser	Pro	Leu	Th <u>r</u> 715	Asp	Ser	Ser	His	Trp 720
Gly	Asn	Lys	Ala	Ile 725											

<210> 27 <211> 7419 <212> DNA <213> Homo Sapiens

<400> 27 60 cggggaccga tccagcctcc ggactctagc ctaggctttt gcaaaaagct atttaggtga 120 cactatagaa ggtacgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc gcageccegt egeeggegga ggegggegeg ggegegtnee etgtggeeag teacceggag 180 gagitggtcg cacaattatg aaagactcgg cttctgctgc tagcgccgga gctgagttag 240 ttctgagaag gtttccctgg gcgttccttg tccggcggcc tctgctgccg cctccggaga 300 360 cgcttcccga tagatggcta caggccgcgg aggaggagga ggtggagttg ctgcccttcc ggagtccgcc ccgtgaggag aatgtcccag aaatcctgga tagaaagcac tttgaccaag 420 480 agggaatgtg tatatattat accaagttcc aaggaccctc acagatgcct tccaggatgt caaatttgtc agcaactcgt caggtgtttt tgtggtcgct tggtcaagca acatgcttgt 540 600 tttactgcaa gtcttgccat gaaatactca gatgtgaaat tgggtgacca ttttaatcag 660 720 gtcataaatt ttcaaggggg ttctcattcc tacagagcta agtatgtgag gctatcatat gacaccaaac ctgaagtcat tctgcaactt ctgcttaaag aatggcaaat ggagttaccc 780 840 aaacttgtta tctctgtaca tgggggcatg cagaaatttg agcttcaccc acgaatcaag 900 cagttgcttg gaaaaggtct tattaaagct gcagttacaa ctggagcctg gattttaact 960 ggaggagtaa acacaggtgt ggcaaaacat gttggagatg ccctcaaaga acatgcttcc 1020 agatcatctc gaaagatttg cactatcgga atagctccat ggggagtgat tgaaaacaga 1080 aatgatcttg ttgggagaga tgtggttgct ccttatcaaa ccttattgaa ccccctgagc 1140 aaattgaatg ttttgaataa tctgcattcc catttcatat tggtggatga tggcactgtt 1200 ggaaagtatg gggcggaagt cagactgaga agagaacttg aaaaaactat taatcagcaa 1260 agaattcatg ctaggattgg ccagggtgtc cctgtggtgg cacttatatt tgagggtggg 1320 ccaaatgtta tcctcacagt tcttgaatac cttcaggaaa gcccccctgt tccagtagtt 1380 gtgtgtaag gaacaggcag agctgcagat ctgctagcgt atattcataa acaaacagaa 1440 gaaggaggga atcttcctga tgcagcagag cccgatatta tttccactat caaaaaaaca 1500 tttaactitg gccagaatga agcacttcat ttatttcaaa cactgatgga gtgcatgaaa 1560 agaaaggagc ttatcactgt tttccatatt gggtcagatg aacatcaaga tatagatgta 1620 gcaatactta ctgcactgct aaaaggtact aatgcatctg catttgacca gcttatcctt acattggcat gggatagagt tgacattgcc aaaaatcatg tatttgttta tggacagcag 1680 tggctggttg gatccttgga acaagctatg cttgatgctc ttgtaatgga tagagttgca 1740 tttgtaaaac ttcttattga aaatggagta agcatgcata aattccttac cattccgaga 1800 1860 ctggaagaac tttacaacac taaacaaggt ccaactaatc caatgctgtt tcatcttgtt 1920 cgagacgtca aacagggaaa tcttcctcca ggatataaga tcactctgat tgatatagga 1980 cttgttattg aatatctcat gggaggaacc tacagatgca cctatactag gaaacgtttt 2040 cgattaatat ataatagtct tggtggaaat aatcggaggt ctggccgaaa tacctccagc 2100 agcactcctc agttgcgaaa gagtcatgaa tcttttggca atagggcaga taaaaaggaa 2160 aaaatgaggc ataaccattt cattaagaca gcacagccct tccgaccaaa gattgataca gttatggaag aaggaaagaa gaaaagaacc aaagatgaaa ttgtagacat tgatgatcca 2220 gaaaccaagc gctttcctta tccacttaat gaacttttaa tttgggcttg ccttatgaag 2280 2340 aggcaggtca tggcccgttt tttatggcaa catggtgaag aatcaatggc taaagcatta 2400 gttgcctgta agatctatcg ttcaatggca tatgaagcaa agcagagtga cctggtagat gatacttcag aagaactaaa acagtattcc aatgattttg gtcagttggc cgttgaatta 2460 2520 ttagaacagt ccttcagaca agatgaaacc atggctatga aattgctcac ttatgaactg aagaactgga gtaattcaac ctgccttaag ttagcagttt cttcaagact tagacctttt 2580 2640 gtagctcaca cctgtacaca aatgttgtta tctgatatgt ggatgggaag gctgaatatg aggaaaaatt cctggtacaa ggtcatacta agcattttag ttccacctgc catattgctg 2700 ttagagtata aaactaaggc tgaaatgtcc catatcccac aatctcaaga tgctcatcag 2760 2820 atgacaatgg atgacagcga aaacaacttt cagaacataa cagaagagat ccccatggaa gtgtttaaag aagtacggat tttggatagt aatgaaggaa agaatgagat ggagatacaa 2880 2940 atgaaatcaa aaaagcttcc aattacgcga aagttttatg ccttttatca tgcaccaatt gtaaaattct ggtttaacac gttggcatat ttaggatttc tgatgcttta tacatttgtg 3000 gttcttgtac aaatggaaca gttaccttca gttcaagaat ggattgttat tgcttatatt 3060 tttacttatg ccattgagaa agtccgtgag atctttatgt ctgaagctgg gaaagtaaac cagaagatta aagtatggtt tagtgattac ttcaacatca gtgatacaat tgccataatt 3120 3180 3240 tctttcttca ttggatttgg actaagattt ggagcaaaat ggaactttgc aaatgcatat gataatcatg tttttgtggc tggaagatta atttactgtc ttaacataat attttggtat 3300 3360 gtgcgtttgc tagattttct agctgtaaat caacaggcag gaccttatgt aatgatgatt ggaaaaatgg tggccaatat gttctacatt gtagtgatta tggctcttgt attacttagt 3420 tttggtgttc ccagaaaggc aatactttat cctcatgaag caccatcttg gactcttgct 3480 aaagatatag tttttcaccc atactggatg atttttggtg aagtttatgc atacgaaatt 3540

-40-

			-4 0-			
gatgtgtgtg	caaatgattc	tgttatccct	caaatctgtg	gtcctgggac	gtggttgact	3600
ccatttcttc	aagcagtcta	cctctttgta	cagtatatca	ttatggttaa	tcttcttatt	3660
gcatttttca	acaatgtgta	tttacaagtg	aaggcaattt	ccaatattgt	atggaagtac	3720
cagcottatc	attttattat	ggcttatcat	gagaaaccag	ttctgcctcc	tccacttatc	3780
attettagee	atatagtttc	tctqttttqc	tgcatatgta	agagaagaaa	gaaagataag	3840
acttccgatg	gaccaaaact	tttcttaaca	gaagaagatc	aaaagaaact	tcatgatttt	3900
gaagaggagt	gtgttgaaat	gtatttcaat	gaaaaagatg	acaaatttca	ttctqqqaqt	3960
gaagagaaga	ttcgtgtcac	ttttgaaaga	gtggaacaga	totocattca	gattaaagaa	4020
gaagagagaa	gtgtcaacta	cataaaaaaa	tcattacaat	cattagattc	tcaaattggc	4080
catttacaaa	atctttcagc	cctgacgga	garacattaa	aaacactcac	tacccagaaa	4140
cattegeaag	ctagcaaagt	testastess	atcacaccac	aactgaggat	ttccaaacac	4200
ttaaataaaa	accttattga	tastaateet	gtaagacctt	ctgtatggaa	aaagcatggt	4260
attatasata	cacttagctc	ctctcttcct	caagacetc	ttgaaagtaa	taatcctttt	4320
gitgiaaata	ttttaatgaa	agatgagaaa	gatccccagt	otaatatatt	taatcaaaac	4380
ttacetaca	taccccagag	agacgacata	asttttccag	aggetgatte	ctcttctaat	4440
ttacctgcag	Laccecagay	ttaaaataa	gaactgcgac	aggetggtte	tagagtagaa	4500
geettattee	caagtgctgt	ccccccca	ttagggatt	catetactac	cataccacat	4560
ctcttaaaaa	tatttaataa	adattadada	agtagaget	catctactag	ttacaaaaaa	4620
ctgtcatccc	caccaaccaa	actititique	agracaccar	ctcagccaag	agaagaaat	4680
cacttggaaa	ctggaaccaa	agattaagaa	accyctiget	teastttaa	agaaggagac	4740
aatacagaat	ttggagcatt	tgtaggacac	agagatagea	Lygatttaca	gaggittaaa	4800
gaaacatcaa	acaagataaa	aatactatcc	aataacaata	cttctgaaaa	Cactttgaaa	4860
cgagtgagtt	ctcttgctgg	atttactgac	tgtcacagaa	cttccattcc	tgttcattca	4920
aaacaagcag	aaaaaatcag	tagaaggcca	tctaccgaag	acactcatga	agtagattcc	
aaagcagctt	taataccgga	ttggttacaa	gatagaccat	caaacagaga	aatgccatct	4980
gaagaaggaa	cattaaatgg	tctcacttct	ccatttaagc	cagctatgga	tacaaattac	5040
tattattcag	ctgtggaaag	aaataacttg	atgaggttat	cacagagcat	tccatttaca	5100
cctgtgcctc	caagagggga	gcctgtcaca	gtgtatcgtt	tggaagagag	ttcacccaac	5160
atactaaata	acagcatgtc	ttcttggtca	caactaggcc	tctgtgccaa	aatagagttt	5220
ttaagcaaag	aggagatggg	aggaggttta	cgaagagctg	tcaaagtaca	gtgtacctgg	5280
tcagaacatg	atatcctcaa	atcagggcat	ctttatatta	tcaaatcttt	tcttccagag	5340
gtggttaata	catggtcaag	tatttataaa	gaagatacag	ttctgcatct	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacgtttg	cctttaatca	aatgaaaccc	5460
aaatccatac	catattctcc	aaggttcctt	gaagttttcc	tgctgtattg	ccattcagca	5520
ggacagtggt	ttgctgtgga	agaatgtatg	actggagaat	ttagaaaata	caacaataat	5580
aatggagatg	agattattcc	aactaatact	ctggaagaga	tcatgctagc	ctttagccac	5640
tggacttacg	aatatacaag	aggggagtta	ctggtacttg	atttgcaagg	tgttggtgaa	5700
aatttgactg	acccatctgt	gataaaagca	gaagaaaaga	gatcctgtga	tatggttttt	5760
ggcccagcaa	atctaggaga	agatgcaatt	aaaaacttca	gagcaaaaca	tcactgtaat	5820
tcttgctgta	gaaagcttaa	acttccagat	ctgaagagga	atgattatac	gcctgataaa	5880
attatatttc	ctcaggatga	gccttcagat	ttgaatcttc	agcctggaaa	ttccaccaaa	5940
gaatcagaat	caactaattc	tgttcgtctg	atgttataat	attaatatta	ctgaatcatt	6000
ggttttgcct	gcacctcaca	gaaatgttac	tgtgtcactt	ttccctcggg	aggaaattgt	6060
ttggtaatat	agaaaggtgt	atgcaagttg	aatttgctga	ctccagcaca	gttaaaaggt	6120
caatattctt	ttgacctgat	taatcagtca	gaaagtccct	ataggataga	gctggcagct	6180
gagaaatttt	aaaggtaatt	gataattagt	atttgtaact	ttttaaaggg	ctctttgtat	6240
agcagaggat	ctcatttgac	tttgttttga	tgagggtgat	gccctctctt	atgtggtaca	6300
ataccattaa	ccaaaggtag	gtgtccatgc	agattttatt	ggcagctgtt	ttattgccat	6360
tcaactaggg	aaatgaagaa	atcacgcagc	cttttggtta	aatggcagtc	aaaattttcc	6420
tcagtgtatt	tagtgtgttc	agtgatgata	tcactggttc	ccaactagat	gcttgttggc	6480
cacqqqaaqq	gaaatgactt	gttctaattc	taggttcaca	gaggtatgag	aagcctgaac	6540
tgaagaccat	tttcaagagg	gacggtattt	atgaatcagg	gttaggctcc	atatttaaag	6600
atagagccag	tttttttt	aaatagaacc	caaattgtgt	aaaaatgtta	attgggtttt	6660
ttaaacattg	ttttatcaag	tcactgttaa	gtagaagaaa	gccatggtaa	actgatacat	6720
aacctaaatt	ataaaaacaa	aaacctaact	cactcgtcaa	gggaagttac	cttttgagga	6780
aagttaaagt	actttttcc	ctatctqtat	ctatagcaac	aacccagaac	ttacaaactt	6840
ctccaaagat	tttattgatt	ottatatcaa	atcagaatgt	aaacatgaac	tcttgcatat	6900
atttaaaatt	gtgttggaac	atttgaacat	gaatgctgtt	tgggtactta	agaaattrat	6960
tcagtnggat	tatcattatg	tganactggc	agattgcagt	gcancettat	gccaataaaa	7020
totaatttar	cagccccaga	tattgttgaa	tattcaacaa	taacaagaaa	agcttttcat	7080
ctaagtttta	tgctttaatt	ttttttttt	ttttttttt	ttcttttatt	tecttagtac	7140
	- 50000000				, , , , , ,	

-41-

taattttaat	ttttatttgg	aagggagcag	tataaagctt	atttgtattt	agtagtgtat	7200
					aatattgatt	7260
gggggtgggg	agaaagaaaa	agtgtattac	ttaaagatac	tatatacskt	ttktatatca	7320
ttaaatcttt	aaaagaaatn	naataaattt	attgtttnca	aaaaaaaac	ccnntaaaaa	7380
aaaaagggcg	gcccctctag	aggatccctc	gaggggccc			7419

<210> 28 <211> 1865 <212> PRT <213> Homo Sapiens

<400> 28 Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys 5 10 Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly 20 25 30 20 25 Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val 40 Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp 55 60 Val Lys Leu Gly Asp His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val 70 75 Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn 85 90 95 85 90 Phe Gln Gly Gly Ser His Ser Tyr Arg Ala Lys Tyr Val Arg Leu Ser 100 105 110 100 105 Tyr Asp Thr Lys Pro Glu Val Ile Leu Gln Leu Leu Leu Lys Glu Trp 120 125 Gln Met Glu Leu Pro Lys Leu Val Ile Ser Val His Gly Gly Met Gln 135 140 Lys Phe Glu Leu His Pro Arg Ile Lys Gln Leu Leu Gly Lys Gly Leu 150 155 Ile Lys Ala Ala Val Thr Thr Gly Ala Trp Ile Leu Thr Gly Gly Val 170 165 175 Asn Thr Gly Val Ala Lys His Val Gly Asp Ala Leu Lys Glu His Ala 190 180 185 Ser Arg Ser Ser Arg Lys Ile Cys Thr Ile Gly Ile Ala Pro Trp Gly 195 200 205 Val Ile Glu Asn Arg Asn Asp Leu Val Gly Arg Asp Val Val Ala Pro 210 220 215 Tyr Gln Thr Leu Leu Asn Pro Leu Ser Lys Leu Asn Val Leu Asn Asn 225 230 235 240 Leu His Ser His Phe Ile Leu Val Asp Asp Gly Thr Val Gly Lys Tyr 255 250 Gly Ala Glu Val Arg Leu Arg Arg Glu Leu Glu Lys Thr Ile Asn Gln 265 270 260 Gln Arg Ile His Ala Arg Ile Gly Gln Gly Val Pro Val Val Ala Leu 275 280 285 280 Ile Phe Glu Gly Gly Pro Asn Val Ile Leu Thr Val Leu Glu Tyr Leu 300 295 Gln Glu Ser Pro Pro Val Pro Val Val Val Cys Glu Gly Thr Gly Arg 305 310 315 Ala Ala Asp Leu Leu Ala Tyr Ile His Lys Gln Thr Glu Glu Gly Gly 325 330 Asn Leu Pro Asp Ala Ala Glu Pro Asp Ile Ile Ser Thr Ile Lys Lys 350 340 345 Thr Phe Asn Phe Gly Gln Asn Glu Ala Leu His Leu Phe Gln Thr Leu 360

WO 00/40614 -42-Met Glu Cys Met Lys Arg Lys Glu Leu Ile Thr Val Phe His Ile Gly Ser Asp Glu His Gln Asp Ile Asp Val Ala Ile Leu Thr Ala Leu Leu Lys Gly Thr Asn Ala Ser Ala Phe Asp Gln Leu Ile Leu Thr Leu Ala Trp Asp Arg Val Asp Ile Ala Lys Asn His Val Phe Val Tyr Gly Gln Gln Trp Leu Val Gly Ser Leu Glu Gln Ala Met Leu Asp Ala Leu Val Met Asp Arg Val Ala Phe Val Lys Leu Leu Ile Glu Asn Gly Val Ser Met His Lys Phe Leu Thr Ile Pro Arg Leu Glu Glu Leu Tyr Asn Thr Lys Gln Gly Pro Thr Asn Pro Met Leu Phe His Leu Val Arg Asp Val Lys Gln Gly Asn Leu Pro Pro Gly Tyr Lys Ile Thr Leu Ile Asp Ile 500 505 510 Gly Leu Val Ile Glu Tyr Leu Met Gly Gly Thr Tyr Arg Cys Thr Tyr Thr Arg Lys Arg Phe Arg Leu Ile Tyr Asn Ser Leu Gly Gly Asn Asn Arg Arg Ser Gly Arg Asn Thr Ser Ser Ser Thr Pro Gln Leu Arg Lys Ser His Glu Ser Phe Gly Asn Arg Ala Asp Lys Lys Glu Lys Met Arg His Asn His Phe Ile Lys Thr Ala Gln Pro Phe Arg Pro Lys Ile Asp Thr Val Met Glu Glu Gly Lys Lys Lys Arg Thr Lys Asp Glu Ile Val Asp Ile Asp Asp Pro Glu Thr Lys Arg Phe Pro Tyr Pro Leu Asn Glu Leu Leu Ile Trp Ala Cys Leu Met Lys Arg Gln Val Met Ala Arg Phe 625 630 635 640 Leu Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met 690 695 700Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr Cys Thr Gln Met Leu Leu Ser Asp Met Trp Met Gly Arg Leu Asn Met Arg Lys Asn Ser Trp Tyr Lys Val Ile Leu Ser Ile Leu Val Pro
755 760 765 Pro Ala Ile Leu Leu Leu Glu Tyr Lys Thr Lys Ala Glu Met Ser His Ile Pro Gln Ser Gln Asp Ala His Gln Met Thr Met Asp Asp Ser Glu Asn Asn Phe Gln Asn Ile Thr Glu Glu Ile Pro Met Glu Val Phe Lys 810 815 Glu Val Arg Ile Leu Asp Ser Asn Glu Gly Lys Asn Glu Met Glu Ile Gln Met Lys Ser Lys Lys Leu Pro Ile Thr Arg Lys Phe Tyr Ala Phe

850	Pro Ile		855					860				
Gly Phe Leu 865		870					875					880
Leu Pro Ser	885					890					895	
Ala Ile Glu	Lys Val 900	Arg	Glu	Ile	Phe 905	Met	Ser	Glu	Ala	Gly 910	Lys	Val
Asn Gln Lys 915	_			920					925			
Thr Ile Ala 930			935					940				
Ala Lys Trp 945		950					955					960
Gly Arg Leu	965					970					975	
Leu Asp Phe	980				985					990		
Ile Gly Lys 995				1000)				1005	5		
Leu Val Leu 1010			1015	5				1020)			
His Glu Ala 1025		1030)				1035	5				1040
Tyr Trp Met	104	5				1050)				1055	5
Ala Asn Asp	1060				1065	5				1070)	
Thr Pro Phe	5			1080)				1085	5		
Val Asn Leu	Leu Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys
1090			1095					1100)			
Ala Ile Ser 1105		1110	Trp	Lys		Gln	Arg 1115	1100 Tyr) His	Phe	Ile	Met 1120
Ala Ile Ser 1105 Ala Tyr His	Glu Lys	1110 Pro 5	Trp) Val	Lys Leu	Pro	Gln Pro 1130	Arg 1115 Pro	1100 Tyr Leu	His	Phe Ile	Ile Leu 1135	Met 1120 Ser
Ala Ile Ser 1105 Ala Tyr His His Ile Val	Glu Lys 112 Ser Leu 1140	1110 Pro 5 Phe	Trp) Val Cys	Lys Leu Cys	Pro Ile 1145	Gln Pro 1130 Cys	Arg 1115 Pro) Lys	1100 Tyr Leu Arg	His His Ile Arg	Phe Ile Lys 1150	Ile Leu 1135 Lys)	Met 1120 Ser 5 Asp
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115	Glu Lys 112 Ser Leu 1140 Asp Gly	1110 Pro 5 Phe Pro	Trp) Val Cys Lys	Lys Leu Cys Leu 1160	Pro Ile 1145 Phe	Gln Pro 1130 Cys Leu	Arg 1115 Pro Lys Thr	1100 Tyr Leu Arg Glu	His Ile Arg Glu 1165	Phe Ile Lys 1150 Asp	Ile Leu 1135 Lys) Gln	Met 1120 Ser Asp Lys
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe	1110 Pro 5 Phe Pro Glu	Trp Val Cys Lys Glu 1175	Lys Leu Cys Leu 1160 Gln	Pro Ile 1145 Phe) Cys	Pro 1130 Cys Leu Val	Arg 1115 Pro) Lys Thr	1100 Tyr Leu Arg Glu Met	His Ile Arg Glu 1165 Tyr	Phe Ile Lys 1150 Asp Phe	Leu 1139 Lys) Gln Asn	Met 1120 Ser Asp Lys
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe	1110 Pro 5 Phe Pro Glu His 1190	Trp Val Cys Lys Glu 1175 Ser	Lys Leu Cys Leu 1160 Gln Gly	Pro Ile 1145 Phe Cys	Pro 1130 Cys Leu Val	Arg 1115 Pro Lys Thr Glu Glu 1195	1100 Tyr Leu Arg Glu Met 1180 Arg	His Ile Arg Glu 1165 Tyr Ile	Phe Ile Lys 1150 Asp Phe Arg	Leu 1135 Lys) Gln Asn	Met 1120 Ser Asp Lys Glu Thr 1200
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120	Pro Pro Glu His 1190 Gln 5	Trp Val Cys Lys Glu 1175 Ser Met	Lys Leu Cys Leu 1160 Gln Gly Cys	Pro Ile 1145 Phe Cys Ser	Pro 1130 Cys Leu Val Glu Gln 1210	Arg 1115 Pro Lys Thr Glu 1195 Ile	1100 Tyr Leu Arg Glu Met 1180 Arg	His Ile Arg Glu 1165 Tyr Ile Glu	Phe Ile Lys 1150 Asp Phe Arg	Leu 1133 Lys Gln Asn Val Gly 1215	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Tyr Ile 1220	Pro Fhe Pro Glu His 1190 Gln 5 Lys	Trp Val Cys Lys Glu 1175 Ser Met	Lys Leu Cys Leu 1160 Gln Gly Cys Ser	Pro Ile 1145 Phe Cys Ser Ile Leu 1225	Gln Pro 1130 Cys Leu Val Glu Gln 1210 Gln	Arg 1115 Pro) Lys Thr Glu 1195 Ile) Ser	1100 Tyr Leu Arg Glu Met 1180 Arg Lys	His Ile Arg Glu 1165 Tyr Ile Glu Asp	Phe Ile Lys 1150 Asp Phe Arg Val Ser	Leu 113: Lys Gln Asn Val Gly 121: Gln	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp 5
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Tyr Ile 1220 Gln Asp	Pro Pro Glu His 1190 Gln Lys Leu	Trp Val Cys Lys Glu 1175 Ser Met Arg	Lys Leu Cys Leu 1160 Gln Gly Cys Ser Ala 1240	Pro Ile 1145 Phe Cys Ser Ile Leu 1225 Leu	Gln Pro 1130 Cys Leu Val Glu Gln 1210 Gln	Arg 1115 Pro Lys Thr Glu 1195 Ile Ser Val	1100 Tyr Leu Arg Glu Met 1180 Lys Lys Leu Asp	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu	Leu 1135 Lys Gln Asn Val Gly 1215 Gln	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp 5 Ile
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123 Leu Thr Ala 1250	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Tyr Ile 1220 Gln Asp 6 Gln Lys	Pro Pro Glu His 1190 Gln Lys Leu Ala	Trp) Val Cys Lys Glu 1175 Ser Met Arg Ser Ser 1255	Lys Leu Cys Leu 1160 Gln Gly Cys Ser Ala 1240	Pro Ile 1145 Phe Cys Ser Ile 1225 Leu Ala	Pro 1130 Cys Leu Val Glu Gln Gln Thr	Arg 1119 Pro Lys Thr Glu 1195 Ile Ser Val	1100 Tyr Leu Arg Glu Met 1180 Arg Lys Leu Asp Val 1260	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr 1245 His	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu Asn	Leu 1135 Lys Gln Asn Val Gly 1215 Gln Lys	Met 1120 Ser Asp Lys Glu Thr 1200 Asp Ile Thr
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123 Leu Thr Ala 1250 Thr Arg Glu 1265	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Tyr Ile 1220 Gln Asp 6 Gln Lys Leu Ser	Pro Pro Glu His 1190 Gln Lys Leu Ala Ile 1270	Trp Val Cys Lys Glu 1175 Ser Met Arg Ser 1255 Ser	Lys Leu Cys Leu 1160 Gln Gly Cys Ser Ala 1240 Glu Lys	Pro Ile 1145 Phe Cys Ser Ile Leu 1225 Leu Ala	Pro 1130 Cys Leu Val Glu Gln Thr Ser	Arg 1119 Pro Lys Thr Glu 1199 Ile Ser Val Lys Ala 1279	1100 Tyr Leu Arg Glu Met 1180 Arg Lys Leu Asp Val 1260 Gln	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr 1245 His	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu Asn Leu	Leu 1135 Lys Gln Asn Val Gly 1215 Gln Lys Glu Ile	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp 5 Ile Thr Lle Asp 1280
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123 Leu Thr Ala 1250 Thr Arg Glu	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Tyr Ile 1220 Gln Asp 6 Gln Lys Leu Ser	Pro Pro Glu His 1190 Gln Lys Leu Ala Ile 1270 Pro	Trp Val Cys Lys Glu 1175 Ser Met Arg Ser 1255 Ser	Lys Leu Cys Leu 1160 Gln Gly Cys Ser Ala 1240 Glu Lys	Pro Ile 1145 Phe Cys Ser Ile Leu 1225 Leu Ala	Pro 1130 Cys Leu Val Glu Gln Thr Ser	Arg 1119 Pro Lys Thr Glu 1195 Ile Ser Val Lys Ala 1275 Lys	1100 Tyr Leu Arg Glu Met 1180 Arg Lys Leu Asp Val 1260 Gln	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr 1245 His	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu Asn Leu	Leu 1135 Lys Gln Asn Val Gly 1215 Gln Lys Glu Ile	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp 5 Ile Thr Ile Asp 1280 Asn
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123 Leu Thr Ala 1250 Thr Arg Glu 1265	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Gln Asp Gln Lys Leu Ser Val Arg 128	Pro Pro Glu His 1190 Gln Lys Leu Ala Ile 1270 Pro 5	Trp Val Cys Lys Glu 1175 Ser Met Arg Ser 1255 Ser	Lys Leu Cys Leu 1160 Gln Gly Cys Ser Ala 1240 Glu Lys Val	Pro Ile 1145 Phe Cys Ser Ile 1225 Leu Ala His	Gln Pro 1130 Cys Leu Val Glu Gln 1210 Gln Thr Ser Leu Lys 1290 Asp	Arg 1119 Pro Lys Thr Glu 1199 Ile Ser Val Lys Ala 1279 Lys	1100 Tyr Leu Arg Glu Met 1180 Arg Lys Leu Asp Val 1260 Gln His	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr 1245 His Asn Gly	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu Asn Leu Val	Leu 1135 Lys Gln Asn Val Gly 1215 Gln Lys Glu Ile Val 1295 Asn	Met 1120 Ser 5 Asp Lys Glu Thr 1200 Asp 5 Ile Thr Le Asp 1280 Asn 5
Ala Ile Ser 1105 Ala Tyr His His Ile Val Lys Thr Ser 115 Lys Leu His 1170 Lys Asp Asp 1185 Phe Glu Arg Arg Val Asn Gly His Leu 123 Leu Thr Ala 1250 Thr Arg Glu 1265 Asp Gly Pro	Glu Lys 112 Ser Leu 1140 Asp Gly 5 Asp Phe Lys Phe Val Glu 120 Gln Asp 6 Gln Lys Leu Ser Val Arg 128 Ser Ser 1300 Asn Ile	Pro Pro Glu His 1190 Gln Lys Leu Ala Ile 1270 Pro Leu	Trp Val Cys Lys Glu 1175 Ser Met Arg Ser 1255 Ser Ser	Lys Leu Cys Leu 1160 Gly Cys Ser Ala 1240 Glu Lys Val	Pro Ile 1145 Phe Cys Ser Ile 1225 Leu Ala His Trp Gly 1305 Asp	Gln Pro 1130 Cys Leu Val Glu Gln 1210 Gln Ser Leu Lys 1290 Asp	Arg 1115 Pro Lys Thr Glu 1195 Ile Ser Val Lys Ala 1275 Lys	Arg Glu Met 1180 Arg Lys Leu Asp Val 1260 Gln His	His Ile Arg Glu 1165 Tyr Ile Glu Asp Thr 1245 His Asn Gly Ser	Phe Ile Lys 1150 Asp Phe Arg Val Ser 1230 Leu Asn Leu Val Asn 1310 Gln	Leu 1135 Lys Gln Asn Val Gly 1215 Glu Ile Val 1295 Asn	Met 1120 Ser Asp Lys Glu Thr 1200 Asp Ile Thr Ile Asp 1280 Asn Fro

Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn 1335 1340 1330 Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val 1350 1355 1345 Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys 1365 1370 1375 Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro 1380 1385 1390 His Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln 1395 1400 1405 Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr 1410 1415 1420 Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe 1435 1430 Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser 1445 1450 1455 Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu 1460 1465 1470 Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser 1475 1480 1485 Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser 1490 1495 1500 Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp 1505 1510 1515 1520 Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly 1525 1530 1535 Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn 1540 1545 1550 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln 1555 1560 1565 Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val 1575 1580 1570 Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser 1590 1595 Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys 1605 1610 1615 Glu Glu Met Gly Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr 1620 1625 1630 Trp Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys 1635 1640 1645 Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu 1650 1655 1660 Asp Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Arg Ala 1665 1670 1675 Ala Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile 1690 1695 1685 Pro Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser 1700 1705 1710 Ala Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg 1715 1720 1725 Lys Tyr Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu 1730 1735 1740 Glu Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg 1750 1755 1760 Gly Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr 1765 1770 1775 Asp Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val 1790 1780 1785 Phe Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala 1800 1805

-45-

<210> 29 <211> 4061 <212> DNA <213> Homo Sapiens

<400> 29

60 qqtctqqaag cagaqccggc ggagggagcg ccggggccct gggctgcagg aggttgcggc 120 ggccgcggca gcatggtggt gccggagaag gagcagagct ggatccccaa gatcttcaag aagaagacct gcacgacgtt catagttgac tccacagatc cgggagggac cttgtgccag 180 240 tqtqqqcqcc cccqqaccqc ccacccqca gtgqccatgq aggatqcctt cggggcagcc gtggtgaccg tgtgggacag cgatgcacac accacggaga agcccaccga tgcctacgga 300 360 qaqctqqact tcacqqqqqc cggccqcaaq cacaqcaatt tcctccqqct ctctqaccqa acggatccag ctgcagttta tagtctggtc acacgcacat ggggcttccg tgccccgaac 420 ctggtggtgt cagtgctggg gggatcgggg ggccccgtcc tccagacctg gctgcaggac 480 ctgctgcgtc gtgggctggt gcgggctgcc cagagcacag gagcctggat tgtcactggg 540 600 qqtctqcaca cqgqcatcqq ccqqcatqtt qqtqtqqctq tacqqqacca tcaqatqqcc agcactgggg gcaccaaggt ggtggccatg ggtgtggccc cctggggtgt ggtccggaat 660 720 agagacacco toateaacco caagggotog ttocotgoga ggtacoggtg gcgcggtgac ccggaggacg gggtccagtt tcccctggac tacaactact cggccttctt cctggtggac 780 840 qacqqcacac acqqctqcct ggggggcgag aaccqcttcc gcttgcgcct ggagtcctac atctcacage agaagaeggg egtgggaggg actggaattg acatecetgt ectgeteete 900 960 ctgattgatg gtgatgagaa gatgttgacg cgaatagaga acgccaccca ggctcagctc ccatgtctcc tcgtggctgg ctcaggggga gctgcggact gcctggcgga gaccctggaa 1020 1080 qacactetqq ceceaqqqaq tqqqqqqqce aggeaaqgeq aageeegaga tegaateaqq cgtttctttc ccaaagggga ccttgaggtc ctgcaggccc aggtggagag gattatgacc 1140 cggaaggagc tcctgacagt ctattcttct gaggatgggt ctgaggaatt cgagaccata 1200 gttttgaagg cccttgtgaa ggcctgtggg agctcggagg cctcagccta cctggatgag 1260 1320 ctgcgtttgg ctgtggcttg gaaccgcgtg gacattgccc agagtgaact ctttcggggg gacatccaat ggcggtcctt ccatctcgaa gcttccctca tggacgccct gctgaatgac 1380 1440 eggeetqaqt teqtqeqett geteatttee caeggeetea geetgggeea etteetgace 1500 ccgatgegee tggeccaact ctacagegeg gegeeeteea actegeteat cegeaacett 1560 ttqqaccaqq cqtcccacaq cqcaqqcacc aaagccccaq ccctaaaagg gggagctgcg gageteegge eccetgaegt ggggcatgtg etgaggatge tgetggggaa gatgtgegeg 1620 1680 ccgaggtace ecteeggggg egeetgggae ecteaceeag geeagggett eggggagage atgtatctgc tctcggacaa ggccacctcg ccgctctcgc tggatgctgg cctcgggcag 1800 qcccctqqa qcqacctqct tctttqqqca ctgttgctga acagggcaca gatggccatg 1860 tacttetggg agatgggtte caatgeagtt tecteagete ttgggggeetg tttgetgete cgggtgatgg cacgcctgga gcctgacgct gaggaggcag cacggaggaa agacctggcg ttcaagtttg aggggatggg cgttgacctc tttggcgagt gctatcgcag cagtgaggtg 1920 1980 2040 aggetque getectet eegtegetge eegetetggg gggatgeeac ttgeeteeag ctggccatgc aagctgacgc ccgtgccttc tttgcccagg atggggtaca gtctctgctg 2100 acacagaagt ggtggggaga tatggccagc actacaccca tctgggccct ggttctcgcc 2160 ttcttttgcc ctccactcat ctacacccgc ctcatcacct tcaggaaatc agaagaggag 2220 cccacacggg aggagctaga gtttgacatg gatagtgtca ttaatgggga agggcctgtc 2280 2340 gggacggcgg acccagccga gaagacgccg ctgggggtcc cgcgccagtc gggccgtccg ggttgctgcg ggggccgctg cggggggcgc cggtgcctac gccgctggtt ccacttctgg 2400 2460 ggcgcgccgg tgaccatctt catgggcaac gtggtcagct acctgctgtt cctgctgctt ttctcgcggg tgctgctcgt ggatttccag ccggcgccgc ccggctccct ggagctgctg 2520 2580 ctctatttct qqqctttcac qctqctqtqc gaggaactqc gccagggcct gagcggaggc gggggcagcc tcgccagcgg gggccccggg cctggccatg cctcactgag ccagcgcctg 2640 egectetace tegeogacag etggaaccag tgegaectag tggeteteac etgetteete 2700 -46-

```
2760
ctgggcgtgg gctgccggct gaccccgggt ttgtaccacc tgggccgcac tgtcctctgc
atcgacttca tggttttcac ggtgcggctg cttcacatct tcacggtcaa caaacagctg
                                                                      2820
gggcccaaga tcgtcatcgt gagcaagatg atgaaggacg tgttcttctt cctcttcttc
                                                                      2880
ctcggcgtgt ggctggtagc ctatggcgtg gccacggagg ggctcctgag gccacgggac
                                                                      2940
agtgacttcc caagtatect gegeegegte ttetacegte cetacetgea gatetteggg
                                                                      3000
cagattcccc aggaggacat ggacgtggcc ctcatggagc acagcaactg ctcgtcggag
                                                                      3060
cccggcttct gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc
                                                                      3120
aactggctgg tggtgctgct cctcgtcatc ttcctgctcg tggccaacat cctgctggtc
                                                                      3180
aacttgctca ttgccatgtt cagttacaca ttcggcaaag tacagggcaa cagcgatctc
                                                                      3240
tactggaagg cgcagcgtta ccgcctcatc cgggaattcc actctcggcc cgcgctggcc
                                                                      3300
ccgcccttta tcgtcatctc ccacttgcgc ctcctgctca ggcaattgtg caggcgaccc
cggagcccc agccgtcctc cccggccctc gagcatttcc gggtttacct ttctaaggaa
                                                                      3420
gccgagcgga agctgctaac gtgggaatcg gtgcataagg agaactttct gctggcacgc
                                                                      3480
gctagggaca agcgggagag cgactccgag cgtctgaagc gcacgtccca gaaggtggac
                                                                      3540
ttggcactga aacagctggg acacatccgc gagtacgaac agcgcctgaa agtgctggag
                                                                      3600
cgggaggtcc agcagtgtag ccgcgtcctg gggtgggtgg ccgaggccct gagccgctct
                                                                      3660
gccttgctgc ccccaggtgg gccgccaccc cctgacctgc ctgggtccaa agactgagcc
                                                                      3720
ctgctggcgg acttcaagga gaagccccca caggggattt tgctcctaga gtaaggctca
                                                                      3780
tctgggcctc ggcccccgca cctggtggcc ttgtccttga ggtgagcccc atgtccatct
                                                                      3840
gggccactgt caggaccacc tttgggagtg tcatccttac aaaccacagc atgcccggct
                                                                      3900
                                                                      3960
cctcccagaa ccagtcccag cctgggagga tcaaggcctg gatcccgggc cgttatccat
ctggaggctg cagggtcctt ggggtaacag ggaccacaga cccctcacca ctcacagatt
                                                                      4020
                                                                      4061
cctcacactg gggaaataaa gccatttcag aggaaaaaaa a
```

<210> 30 <211> 1214 <212> PRT

<213> Homo Sapiens

<400> 30 Met Val Val Pro Glu Lys Glu Gln Ser Trp Ile Pro Lys Ile Phe Lys 10 Lys Lys Thr Cys Thr Thr Phe Ile Val Asp Ser Thr Asp Pro Gly Gly 25 30 20 Thr Leu Cys Gln Cys Gly Arg Pro Arg Thr Ala His Pro Ala Val Ala 40 Met Glu Asp Ala Phe Gly Ala Ala Val Val Thr Val Trp Asp Ser Asp 55 Ala His Thr Thr Glu Lys Pro Thr Asp Ala Tyr Gly Glu Leu Asp Phe 65 70 75 80 Thr Gly Ala Gly Arg Lys His Ser Asn Phe Leu Arg Leu Ser Asp Arg 90 85 Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe 105 100 Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro 125 120 Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg 140 135 Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr 155 150 Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala 170 165 Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly 190 185 Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro 200 205 Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro 215 220 Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His 235

-47-Gly Cys Leu Gly Gly Glu Asn Arg Phe Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys Thr Gly Val Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Ile Asp Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln Ala Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu Ala Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe Arg Gly 405 410 Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu Met Asp Ala Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu Ile Ser His Gly Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly Thr Lys Ala Pro Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro Asp Val Gly His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro Arg Tyr Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe Gly Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala 530 535 Thr Ser Pro Leu Ser Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Trp Ala Leu Leu Asn Arg Ala Gln Met Ala Met 570 . 575 Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln 645 650 655 Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val 665 670 Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu

-48-

				725					730					Pro 735	
_			740					745					750	Arg	
		755					760					765		Arg	
	770					775					780			Phe	
785					790					795				Arg	800
			_	805					810	•				Leu 815	
			820					825					830	Gln	
		835					840					845		Pro	
	850					855					860			Ser	
865		_	_		870					875				Val	880
_				885					890					Leu 895	
	•		900					905					910	Thr	
	_	915					920					925		Met	
_	930					935					940			Ala	
945					950					955				Phe	960
				965					970					Phe 975	
			980					985					990	Ser	
		995					1000)				100	5	Gln	
_	1010)				1015	5				1020)		Leu	
102	5				1030)				103	5			Leu	1040
				1045	5				105	0				Asp 1059	5
			1060)				1065	5				107		
		1075	5				1080)				108	5	Leu	
	1090)				1099	5				1100)		Ser	
110	5				1110)				111	5			Arg	1120
				1125	5				113	0				Ala 113	5
	_		1140)				1145	5				115		
		115	5				1160)				116	5	Glu	
	1170)				1179	5				1180	0		Ser	
Val 118		Gly	Trp	Val	Ala 1190		Ala	Leu	Ser	Arg 119		Ala	Leu	Leu	Pro 1200

-49-

Pro Gly Gly Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp 1210

> <210> 31 <211> 4646 <212> DNA <213> Homo Sapiens

<400> 31

tegacecacg egteegeeca egegteegee caegegteeg eccaegegte egeceaegeg 60 teegeecaeg egteeggggt gaaagmramy emygektsms aaaaaeegte aettaggaaa 120 agatgteett tegggeagee aggeteagea tgaggaacag aaggaatgae actetggaea 180 240 gcacceggae ectgtactee agegegtete ggageacaga ettgtettae agtgaaageg acttggtgaa ttttattcaa gcaaatttta agaaacgaga atgtgtcttc tttaccaaag 300 360 attccaaggc cacggagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag gcacccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaatttecta 420 480 ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggaag tatatacgtc tgteetgega caeggaegeg gaaateettt aegagetget gaeeeageae tggeaeetga 540 aaacacccaa cctggtcatt tctgtgaccg ggggcgccaa gaacttcgcc ctgaagccgc 600 gcatgcgcaa gatcttcagc cggctcatct acatcgcgca gtccaaaggt gcttggattc 660 720 teacgggagg cacccattat ggcctgatga agtacatcgg ggaggtggtg agagataaca ccatcagcag gagttcagag gagaatattg tggccattgg catagcagct tggggcatgg 780 840 tetecaaccg ggacaccete atcaggaatt gcgatgctga gggctatttt ttagcccagt acettatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt 900 tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaag ctccggaatc 960 agctagagaa gtatatetet gagegeacta tteaagatte caactatggt ggeaagatee 1020 ccattgtgtg ttttgcccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca 1080 tcaaaaataa aatteettgt gtggtggtgg aaggeteggg ccagateget gatgtgateg 1140 ctagectggt ggaggtggag gatgeeetga catettetge egteaaggag aagetggtge 1200 getttttace cegeacggtg teceggetge etgaggagga gaetgagagt tggateaaat 1260 ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaaaatg gaagaagctg 1320 gggatgaaat tgtgagcaat gccatctcct acgctctata caaagccttc agcaccagtg 1380 1440 agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg acttagecaa tgatgagatt tteaceaatg accgeegatg ggagtetget gacetteaag 1500 aagtcatgtt tacggctctc ataaaggaca gacccaagtt tgtccgcctc tttctggaga 1560 atggettgaa cetacggaag tttetcacce atgatgteet caetgaacte ttetccaace 1620 1680 acttcagcac gettgtgtac eggaatetge agategecaa gaatteetat aatgatgeee tectcaegtt tgtetggaaa etggttgega aetteegaag aggetteegg aaggaagaea 1740 1800 gaaatggccg ggacgagatg gacatagaac tecacgacgt gteteetatt acteggcace 1860 ccctgcaagc tctcttcatc tgggccattc ttcagaataa gaaggaactc tccaaagtca tttgggagca gaccaggggc tgcactctgg cagccctggg agccagcaag cttctgaaga 1920 ctetggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaatg 1980 agtacgagac ccgggctgtt gagctgttca ctgagtgtta cagcagcgat gaagacttgg 2040 cagaacaget getggtetat teetgtgaag ettggggtgg aageaactgt etggagetgg 2100 cggtggaggc cacagaccag catticatcg cccagcctgg ggtccagaat tttctttcta 2160 agcaatggta tggagagatt tecegagaca ceaagaactg gaagattate etgtgtetgt 2220 2280 ttattatacc cttggtgggc tgtggctttg tatcatttag gaagaaacct gtcgacaagc acaagaaget getttggtae tatgtggegt tettcacete cecettegtg gtetteteet 2340 ggaatgtggt cttctacatc gccttcctcc tgctgtttgc ctacgtgctg ctcatggatt 2400 tecatteggt gecacacece ecegagetgg teetgtacte getggtettt gteetettet 2460 gtgatgaagt gagacagtgg tacgtaaatg gggtgaatta ttttactgac ctgtggaatg 2520 tgatggacac gctggggctt ttttacttca tagcaggaat tgtatttcgg ctccactctt 2580 ctaataaaag ctctttgtat tctggacgag tcattttctg tctggactac attatttca 2640 ctctaagatt gatccacatt tttactgtaa gcagaaactt aggacccaag attataatgc 2700 tgcagaggat gctgatcgat gtgttcttct tcctgttcct ctttgcggtg tggatggtgg 2760 2820 cctttggcgt ggccaggcaa gggatcctta ggcagaatga gcagcgctgg aggtggatat teegtteggt catetacgag ecetacetgg ceatgttegg ecaggtgeec agtgacgtgg 2880 2940 atggtaccac gtatgacttt gcccactgca ccttcactgg gaatgagtcc aagccactgt gtgtggaget ggatgageac aacctgeece ggtteecega gtggateace ateceectgg 3000 3060 tgtgcatcta catgttatcc accaacatcc tgctggtcaa cctgctggtc gccatgtttg

-50-

```
gctacacggt gggcaccgtc caggagaaca atgaccaggt ctggaagttc cagaggtact
                                                                      3120
tectggtgca ggagtactgc agecgeetea atateceett eccetteate gtettegett
acttctacat ggtggtgaag aagtgcttca agtgttgctg caaggagaaa aacatggagt
                                                                      3240
cttctgtctg ctgtttcaaa aatgaagaca atgagactct ggcatgggag ggtgtcatga
                                                                      3300
aggaaaacta ccttgtcaag atcaacacaa aagccaacga cacctcagag gaaatgaggc
                                                                      3360
atcgatttag acaactggat acaaagctta atgatctcaa gggtcttctg aaagagattg
                                                                      3420
ctaataaaat caaataaaac tgtatgaact ctaatggaga aaaatctaat tatagcaaga
                                                                      3480
tcatattaag gaatgctgat gaacaatttt gctatcgact actaaatgag agattttcag
                                                                      3540
acccctgggt acatggtgga tgattttaaa tcaccctagt gtgctgagac cttgagaata
aagtgtgtga ttggtttcat acttgaagac ggatataaag gaagaatatt tcctttatgt
                                                                      3660
gtttctccag aatggtgcct gtttctctct gtgtctcaat gcctgggact ggaggttgat
                                                                      3720
agtttaagtg tgttcttacc gcctcctttt tcctttaatc ttatttttga tgaacacata
                                                                      3780
tataggagaa catctatcct atgaataaga acctggtcat gctttactcc tgtattgtta
                                                                      3840
ttttgttcat ttccaattga ttctctactt ttcccttttt tgtattatgt gactaattag
                                                                      3900
                                                                      3960
ttggcatatt gtwaaaagtc tctcaaatta ggccagattc taaaacatgc tgcagcaaga
ggaccccgct ctcttcagga aaagtgtttt catttctcag gatgcttctt acctgtcaga
                                                                      4020
                                                                      4080
ggaggtgaca aggcagtete ttgetetett ggaeteacea ggeteetatt gaaggaacea
ccccattcc taaatatgtg aaaagtcgcc caaaatgcaa ccttgaaagg cactactgac
                                                                      4140
                                                                      4200
tttgttctta ttggatactc ctcttattta ttattttcc attaaaaata atagctggct
attatagaaa atttagacca tacagagatg tagaaagaac ataaattgtc cccattacct
                                                                      4260
                                                                      4320
taaggtaatc actgctaaca atttctggat ggtttttcaa gtctattttt tttctatgta
tgtctcaatt ctctttcaaa attttacaga atgttatcat actacatata tacttttat
                                                                      4380
                                                                      4440
gtaagctttt tcacttagta ttttatcaaa tatgttttta ttatattcat agccttctta
aacattatat caataattgc ataataggca acetetageg attaccataa tittgeteat
                                                                      4500
                                                                      4560
tgaaggctat ctccagttga tcattgggat gagcatcttt gtgcatgaat cctattgctg
tatttgggaa aattttccaa ggttagattc caataaatat ctatttatta ttaaaaaaaa
                                                                      4620
                                                                      4646
aaaaaaagg gcggccgctc tagagt
```

<210> 32

<211> 1104

<212> PRT

<213> Homo Sapiens

<400> 32

Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg Arg Asn Asp 10 Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser Arg Ser Thr 25 Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile Gln Ala Asn 40 Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser Lys Ala Thr 55 Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly 75 Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys 90 Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly 105 110 100 Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile 125 120 Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu 140 135 Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro Arg 155 150 Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser Lys Gly 170 165 Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met Lys Tyr Ile 185 190 180 Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser Glu Glu Asn -51-

											_		_	_	_
	Val 210					215					220				
Thr 225	Leu	Ile	Arg	Asn	Cys 230	Asp	Ala	Glu	Gly	Tyr 235	Phe	Leu	Ala	Gln	Tyr 240
	Met	Asp	Asp	Phe 245		Arg	Asp	Pro	Leu 250	Cys	Ile	Leu	Asp	Asn 255	Asn
His	Thr	His	Leu 260		Leu	Val	Asp	Asn 265		Cys	His	Gly	His 270	Pro	Thr
Val	Glu	Ala 275		Leu	Arg	Asn	Gln 280		Glu	Lys	Tyr	Ile 285	_	Glu	Arg
Thr	Ile 290		Asp	Ser	Asn	Tyr 295		Gly	Lys	Ile	Pro 300		Val	Cys	Phe
Ala 305	Gln	Gly	Gly	Gly	Lys 310		Thr	Leu	Lys	Ala 315		Asn	Thr	Ser	Ile 320
Lys	Asn	Lys	Ile	Pro 325		Val	Val	Val	Glu 330		Ser	Gly	Gln	Ile 335	Ala
Asp	Val	Ile	Ala 340		Leu	Val	Glu	Val 345		Asp	Ala	Leu	Thr 350	Ser	Ser
Ala	Val	Lys 355		Lys	Leu	Val	Arg 360		Leu	Pro	Arg	Thr 365	Val	Ser	Arg
Leu	Pro 370		Glu	Glu	Thr	Glu 375	Ser	Trp	Ile	Lys	Trp 380	Leu	Lys	Glu	Ile
Leu 385	Glu	Cys	Ser	His	Leu 390		Thr	Val	Ile	Lys 395	Met	Glu	Glu	Ala	Gly 400
Asp	Glu	Ile	Val	Ser 405	Asn	Ala	Ile	Ser	Tyr 410	Ala	Leu	Tyr	Lys	Ala 415	Phe
Ser	Thr	Ser	Glu 420	Gln	Asp	Lys	Asp	Asn 425	Trp	Asn	Gly	Gln	Leu 430	Lys	Leu
	Leu	435					440					445			
	Asp 450					455					460				
465	Leu				470					475					480
	Leu			485					490					495	
	Ser		500					505					510		
	Asn	515					520					525			
	Asn 530					535					540				
545	Met				550					555					560
	Gln			565					570					575	
	Lys		580					585					590		
	Ala	595					600					605			
	Ala 610					615					620				
625	Val				630					635					640
	Gln			645					650					655	
	Glu		660					665					670		
Gly	Val	Gln 675	Asn	Phe	Leu	Ser	Lys 680	Gln	Trp	Tyr	Gly	Glu 685	Ile	Ser	Arg

Asp	Thr 690	Lys	Asn	Trp	Lys	Ile 695	lle	Leu	Cys	Leu	Phe 700	Ile	Ile	Pro	Leu
705	Gly		Gly		710					715					720
			Leu	725					730					735	
			Trp 740					745					750		
		755	Leu				760					765			
	770		Tyr			775					780				
785	_		Val		790					795					800
			Leu	805					810					815	
			Ser 820					825					830		
_		835	Tyr				840					845			
	850		Asn			855					860				
865			Phe		870					875					880
			Ala	885					890					895	
			Phe 900					905					910		
		915	Pro				920					925			
	930		Thr			935					940				
945			Leu		950					955					960
			Met	965					970					975	
			Gly 980					985					990		
		995	Phe				1000	כ				100	5		
	1010)	Pro			1015	5				1020	3			
1025	5		Суѕ		1030)				103	5				104
			Cys	1045	5				1050	ס				105	5
			Lys 1060)				106	5				107	0	
_		1079					1080)				108	5		
Leu	Asn 1090		Leu	Lys	Gly	Leu 109!		Lys	Glu	ıle	Ala 110	Asn 0	гуѕ	тте	гÀs

INTERNATIONAL SEARCH REPORT

In' ritional Application No PCT/US 99/29996

A. CLASSI IPC 7	FICATION OF SUBJECT MATTER CO7K14/705 C12N15/12 C12Q1/68 G01N33/53 A61K38/17	8 C12N5/10 C07K	16/28				
According to	o International Patent Classification (IPC) or to both national classific	cation and IPC					
	SEARCHED						
Minimum do IPC 7	ocumentation searched (classification system followed by classification C12N C07K C12Q A61K G01N	ion symbols)					
Documentat	tion searched other than minimum documentation to the extent that	such documents are included in the fields s	earched				
Electronic d	ata base consulted during the international search (name of data ba	ase and, where practical, search terms use	d)				
BIOSIS	, EPO-Internal, WPI Data, PAJ, MEDL HEM ABS Data, STRAND, GENSEQ, EMBL						
C. DOCUMI	ENTS CONSIDERED TO BE RELEVANT						
Category *	Citation of document, with indication, where appropriate, of the re	levant passages	Relevant to claim No.				
x	DATABASE GENEMBL 'Online! 16 February 1998 (1998-02-16) STRAUSBERG,R.: "ob70f05.s1 NCI_CO Homo sapiens cDNA clone IMAGE:13: mRNA sequence" XP002138823 Accession AA809355	GAP_GCB1 36737 3',	1,2, 6-19, 25-35				
X	DATABASE GENEMBL 'Online! 10 July 1998 (1998-07-10) MARRA ET AL.: "ub28d10.r1 Soares musculus cDNA clone IMAGE:1379059 sequence" XP002149803 Accession AI050262	1,6-19, 25-35					
	<u>-</u>	-/ -					
X Furth	ner documents are listed in the continuation of box C.	X Patent family members are listed	I in annex.				
"A" docume	* Special categories of cited documents : "T" later document published after the international filling date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the						
"E" earlier d	considered to be of particular relevance Invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention						
"L" docume which	filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "Cannot be considered novel or cannot be considered novel or can						
"O" docume	"O" document referring to an oral disclosure, use, exhibition or document is combined with one or more other such document of the combination being obvious to a person skilled						
P docume	other means """ "Occurrent published prior to the international filing date but later than the priority date claimed """ "Occurrent means and the same patent family """ "Occurrent means and the same patent family						
Date of the	Date of the actual completion of the international search Date of mailing of the international search report						
10	6 October 2000	3 0. 1	0. (10				
Name and m	nailing address of the ISA	Authorized officer					
	European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk						
	Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG, A					

INTERNATIONAL SEARCH REPORT

In Itional Application No PCT/US 99/29996

Continue	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	PC1/US 99/29996
ategory *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 19 July 1997 (1997-07-19) STRAUSBERG, R.: "ni64e11.s1 NCI_CGAP_Pr12 Homo sapiens cDNA clone IMAGE:981644 mRNA sequence" XP002148641 Accession AA523749	1,3, 10-19, 25-35
x	 WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16)	1,4, 6-19, 25-35
Y	page 4, line 7 -page 5, line 13 page 5, line 24 -page 7, line 28 SEQ ID NOs. 9 and 25	20-24
X	WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27)	1,4, 6-19, 25-35
Y	page 7, paragraph 2 page 9, paragraphs 2,3 page 13 -page 17 page 21, paragraph 3 SEQ ID NOs: 109 and 112	20-24
x	DATABASE GENEMBL 'Online! 18 November 1997 (1997-11-18) STRAUSBERG, R.: "nt76b07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204405, mRNA"	1,5-19, 25-35
Y	XP002148642 Accession AA654650	20-24
Y	DATABASE GENEMBL 'Online! 30 November 1998 (1998-11-30) SHIMIZU, N.: "Homo sapiens mRNA complete cds." XP002148643 Accession number AB001535 -& NAGAMINE ET AL.: "Molecular cloning of a novel putative Ca2+ channel protein (TRPC7) highly expressed in brain" GENOMICS, vol. 54, 15 November 1998 (1998-11-15), pages 124-131, XP000938744 the whole document	20-24
	-/	

INTERNATIONAL SEARCH REPORT

Int dional Application No PCT/US 99/29996

	TO THE PROPERTY OF THE PROPERT	PCT/US 99/29996
Category °	ation) DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
JaileyOry *	Ondoon or occurrent, with marcoon, where appropriate, or the relevant possages	
A	ZHU, XI ET AL: "Molecular cloning of a widely expressed human homologue for the Drosophila trp gene." FEBS LETTERS, (1995) VOL. 373, NO. 3, PP. 193-198., XP000907241 page 194; figures 1,3	20,21, 23,25, 26,28, 29,31
A	HUNTER JOHN J ET AL: "Chromosomal localization and genomic characterization of the mouse melastatin gene (MIsn1)." GENOMICS NOV. 15, 1998, vol. 54, no. 1, 15 November 1998 (1998-11-15), pages 116-123, XP000910696 ISSN: 0888-7543 cited in the application page 119; figure 2	20,21,23
A	WES PAUL D ET AL: "TRPC1, a human homolog of a Drosophila store-operated channel." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 1995, vol. 92, no. 21, 1995, pages 9652-9656, XP002138820 ISSN: 0027-8424 the whole document	20,21, 23,25, 26,28, 29,31
A	ZHU, XI ET AL: "Trp, A novel mammalian gene family essential for agonist-activated capacitative Ca-2+ entry." CELL, vol. 85, no. 5, 1996, pages 661-671, XP000907242 page 662 page 665 figures 1,5,6	20,21, 25,26, 28,29,31
A	GARCIA REYNALDO L ET AL: "Differential expression of mammalian TRP homologues across tissues and cell lines." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1997, vol. 239, no. 1, 1997, pages 279–283, XP002138822 ISSN: 0006-291X See Materials and Methods figure 1	25,26, 28-30

INTERNATIONAL SEARCH REPORT

In ational Application No PCT/US 99/29996

		PCT/US 99/29996				
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT Category Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No.						
Category *	CHARDON OF OUCUMENT, WITH INDICATION, WHERE APPROPRIATE, OF the relevant passages	naovani w danii 140.				
A	SINKINS WILLIAM G ET AL: "Functional expression of TrpC1: A human homologue of the Drosophila Trp channel." BIOCHEMICAL JOURNAL APRIL, 1998, vol. 331, no. 1, April 1998 (1998-04), pages 331-339, XP000864583 ISSN: 0264-6021 page 333-335; figures 3-5	24				
A	PREUSS KLAUS-DIETER ET AL: "Expression and characterization of a trp1 homolog from rat." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS NOV. 7, 1997, vol. 240, no. 1, 7 November 1997 (1997-11-07), pages 167-172, XP002138821 ISSN: 0006-291X figure 2	24				
A	OBUKHOV, ALEXANDER G. ET AL: "Direct activation of trpl cation channels by G-alpha-11 subunits." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, (1996) VOL. 15, NO. 21, PP. 5833-5838., XP000907243 figure 2	24				
P,X	WO 99 09199 A (RYAZANOV ALEXEY G ;PAVUR KAREN S (US); HAIT WILLIAM N (US); UNIV M) 25 February 1999 (1999-02-25) see melanome kinase polynucleotide and polypeptide sequences on page 16-17	1,3, 10-19, 25-36				
Ρ,Χ	WO 99 09166 A (SHAPERO MICHAEL H ;DENDREON CORP (US); LAUS REINER (US); TSAVALER) 25 February 1999 (1999-02-25) page 17, line 24 -page 18, line 9 page 25, line 19-32 page 28, line 1-4 SEQ ID NOS: 27, 28 and 31.	1,5-19, 25-35				
Т	SCHARENBERG A M ET AL: "MLSN-1/SOC-1 defines a widely expressed Ca2+/cation channel family involved in Ca2+ homeostasis and store-operated Ca2+ signaling." FIFTY-THIRD ANNUAL MEETING OF THE SOCIETY OF GENERAL PHYSIOLOGISTS; WOODS HOLE, MASSACHUSETTS, USA; SEPTEMBER 9-11, 1999, vol. 114, no. 1, July 1999 (1999-07), page 14a XP000910708 Journal of General Physiology July, 1999 ISSN: 0022-1295					

..temational application No. PCT/US 99/29996

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)							
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:							
Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:							
2. X Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210							
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).							
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)							
This International Searching Authority found multiple inventions in this international application, as follows:							
see additional sheet							
As a result of the prior review under R. 40.2(e) PCT, no additional fees are to be refunded.							
1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.							
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.							
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: 1-36							
4. No required additional search fees were timely paid by the applicant. Consequently, this international Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:							
Remark on Protest X The additional search fees were accompanied by the applicant's protest. No protest accompanied the payment of additional search fees.							

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 1-5, 10-13, 16-19, 32-35 relate to an extremely large number of possible polynucleotides, polypeptides encoded by them, binding polypeptides, and kits and pharmaceutical compositions containing said polypeptides and polynucleotides. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the polynucleotide of SEQ ID NOs: 1, 27, 29 and 31 and the corresponding polypeptide of SEQ ID NOs: 2, 28, 30 and 32.

Present claims 16 and 17 relate to an extremely large number of possible compounds, namely, a polypeptide that binds to the polypeptide of the invention. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to an antibody, antibody fragment, F(ab)2 fragment or a fragment including a CDR3 region selective for the polypeptides of the invention.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1, 6-36 (partially) and 2 (complete)

An isolated nucleic acid molecule comprising a nucleic acid molecule that hybridizes to a nucleic acid molecule of SEQ ID NO:1 and which code for a SOC/CRAC polypeptide, nucleic acid molecules that differ in codon sequence due to degeneracy of the genetic code and complement thereof, polynucleotides which are not identical to the SEQ ID or sequences of GenBank accession number of Table 1; expression vector, host cells; polypeptide encoded thereof (SEQ ID NO:2); polypeptides binding to the polypeptide of SEQ ID NO:2, including antibodies; kits comprising agents that selectively bind to the polynucleotide (SEQ ID NO:1) or polypeptide (SEQ ID NO:2) of the invention; pharmaceutical compositions containing the polynucleotide or polypeptides of the invention; a method for isolating the SOC/CRAC molecule having SOC/CRAC calcium channel activity comprising contacting a binding molecule that is SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing SOC/CRAC molecules allowing the formation of the complex, detecting the formation of the complex, isolating the SOC/CRAC molecule and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity; a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity; a method to determine the level of SOC/CRAC expression in a subject, including expression of SOC/CRAC polypeptide or mRNA in a tissue or biological fluid sample using PCR, Northern blotting, and mono— and polyclonal antisera and a method for identifying agents useful in the modulation of the SOC/CRAC polypeptide kinase activity, comprising the use of aminoacids 999-1180 from SEQ ID NO:4 as a candidate kinase.

2. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:3 and to the encoded polypeptide of SEQ ID NO:4

3. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:5 and to the encoded polypeptide of SEQ ID NO:6

4. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:7 and to the encoded polypeptide of SEQ ID NO:8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1,6-36 (partially) and 37 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:23 and to the encoded polypeptide of SEQ ID NO:24 $\,$

6. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:25 and to the encoded polypeptide of SEQ ID NO:26 $\,$

7. Claims: 1,10-36 (partially) and 3 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:27 and to the encoded polypeptide of SEQ ID NO:28 $\,$

8. Claims: 1,6-36 (partially) and 4 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:29 and to the encoded polypeptide of SEQ ID NO:30 $\,$

9. Claims: 1,6-36 (partially) and 5 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:31 and to the encoded polypeptide of SEQ ID NO:32.

INTERNATIONAL SEARCH REPORT

Information on patent family members

In Itional Application No
PCT/US 99/29996

Patent document cited in search repo	rt	Publication date	Patent family member(s)		Publication date
WO 9815657	A	16-04-1998	US EP US	5919638 A 0954599 A 6110675 A	06-07-1999 10-11-1999 29-08-2000
WO 9837093	A	27-08-1998	AU CN EP NO PL ZA	6181898 A 1252837 T 1005546 A 994069 A 335348 A 9801585 A	09-09-1998 10-05-2000 07-06-2000 22-10-1999 25-04-2000 04-09-1998
WO 9909199	Α	25-02-1999	AU	9110098 A	08-03-1999
WO 9909166	A	25-02-1999	AU EP	9021898 A 1005549 A	08-03-1999 07-06-2000

CORRECTED VERSION

(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 13 July 2000 (13.07.2000)

PCT

(10) International Publication Number WO 00/40614 A3

- (51) International Patent Classification7: C07K 14/705, C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N 33/53, A61K 38/17
- (21) International Application Number: PCT/US99/29996
- (22) International Filing Date:

20 December 1999 (20.12.1999)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/114,220 30 December 1998 (30.12.1998) US 60/120,018 29 January 1999 (29.01.1999) US 60/140,415 22 June 1999 (22.06.1999) US

- (71) Applicant (for all designated States except US): BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 1 Deaconess Road, Boston, MA 02215 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): SCHARENBERG, Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA 02420 (US).

- (74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).
- (81) Designated States (national): AU, CA, JP, US.
- (84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

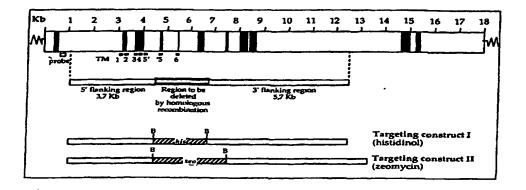
Published:

with international search report

- (88) Date of publication of the international search report: 22 February 2001
- (48) Date of publication of this corrected version: 30 August 2001
- (15) Information about Correction: see PCT Gazette No. 35/2001 of 30 August 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CHARACTERIZATION OF THE SOC/CRAC CALCIUM CHANNEL PROTEIN FAMILY



(57) Abstract: Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.

00/40614 A3

CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

Field of the Invention

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

Background of the Invention

5

10

15

20

25

30

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca²⁺ ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca²⁺ ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca²⁺ cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed store operated calcium influx, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

-2-Summary of the Invention

5

10

15

20

25

30

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca2+ flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca2+ from intracellular calcium stores, allowing Ca2+ influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25 ·

30

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

- a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;

5

10

15

20

25

30

- c) isolating the SOC/CRAC molecule from the complex; and
- d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

5

10

15

20

25

30

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)2, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules in vivo or in vitro.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;
- b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

5

10

15

20

25

30

- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

- a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and
- b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

5

10

15

20

25

30

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, *inter alia*, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEQ ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

5

10

15

20

25

30

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a C. Elegans polypeptide at the c05c12.3 locus.

SEQ ID NO:14 is the amino acid sequence of a *C. Elegans* polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a C. Elegans polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

WO 00/40614

5

10

15

20

25

30

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEQ ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEQ ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

PCT/US99/29996

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEQ ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEQ ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

<u>Figure 1</u> is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca²⁺ flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca²⁺ from intracellular calcium stores, allowing Ca²⁺ influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

5

10

15

20.

25

30

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca²⁺ fluxing and, in particular, lymphocyte Ca²⁺ fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

5

10

15

20

25

30

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid" refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca²⁺ transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

-12-

Science, 1992; 256:498-502; and Hoth. M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

- (a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).

5

10

15

20

25

30

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

5

10

15

20

25

30

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified in vitro by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

5

10

15

20

25

30

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEQ ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEQ ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

-15-

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

5

10

15

20

25

30

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (ftp:/ncbi.nlm.nih.gov/pub/). Exemplary tools include the BLAST system available at http://www.ncbi.nlm.nih.gov. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVetor sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphoimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC cDNA clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, in vitro or in vivo, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

5

10

15

20

25

30

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

-17-

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

5

10

15

20

25

30

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

5

10

15

20

25

30

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., Nat. Med. 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In

-19-

addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5,SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue experimentation.

5

10

15

20

25

30

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, alkylphosphonothioates, phosphoramidates, phosphorodithioates, phosphate esters, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

-20-

5

10

15

20

25

30

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-Oalkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as <u>E.coli</u> and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

-21-

5

10

15

20

25

30

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed -22-

5

10

15

20

25

30

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the group consisting of SEQ ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

WO 00/40614 PCT/US99/29996 -23-

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

5

10

15

20

25

30

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1α, which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., <u>E. coli</u>), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

"knock-outs" in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

5

10

15

20

25

30

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca²⁺ fluxing, high selectivity, a unitary

-25-

5

10

15

20

25

30

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in Science 278:82-87, 1997, whereby proteins can be designed de novo. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

10

15

20

25

30

PCT/US99/29996 WO 00/40614

Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., E. coli, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. Molecular Cloning: A Laboratory Manual, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or Current Protocols in Molecular Biology, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

WO 00/40614 PCT/US99/29996 -27-

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

5

10

15

20

25

30

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described First, the invention permits isolation of SOC/CRAC polypeptides, elsewhere herein. including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

WO 00/40614 PCT/US99/29996
-28-

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

5

10

15

20

25

30

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

5

10

15

20

25

30

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by

the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

5

10

15

20

25

30

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

WO 00/40614 PCT/US99/29996 -31-

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

5

10

15

20

25

30

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

WO 00/40614 PCT/US99/29996 -32-

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

5

10

15

20

25

30

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention. a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

5

10

15

20

25

30

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. Trans-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such trans-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic nonhuman animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

5

10

15

20

25

30

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca²⁺ fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled in vitro protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

5

10

15

20

25

30

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca²⁺. Such separation is preferred so that a change in Ca²⁺ concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cell-membrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in <u>Capacitative Calcium Entry</u>, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel activation and a given reference calcium concentration. Detection of a decrease in the

WO 00/40614 PCT/US99/29996 -36-

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5

10

15

20

25

30

SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or

polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

5

10

15

20

25

30

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include dihydropyridines (DHPs), phenylalkylamines, omega such as (omega.-CgTx) and pyrazonoylguanidines.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

5

10

15

20

25

30

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromotograpic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β-galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

5

10

15

20

25

30

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEO ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and Dictyostelium myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., Proc Natl Acad Sci U S A, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or Dictyostelium MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

-40-

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10⁷ M⁻¹, more preferably at least about 10⁸ M⁻¹, and most preferably at least about 10⁹ M⁻¹. The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

5

10

15

20

25

30

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced in vitro or in vivo in a host. Such techniques include transfection of nucleic acid-CaPO₄ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

-41-

polyhydroxybutyric polyanhydrides. polyesteramides, polyorthoesters, acid, and Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sylastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

5

10

15

20

25

30

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, are used herein, means that the implant is constructed and arranged to delivery therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing ex vivo gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction in vitro of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. In vivo gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

5

10

15

20

25

30

As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode C. elegans is nearing completion, and IP3-dependent calcium signals have recently been shown to be required for one or more aspects of C. elegans development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily excluding all proteins below approximately 1000 amino acids in size, followed by focusing on remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated CO5C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was noteable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

We then used C05C12.3 for BLAST alignment screening of the rest of the *C. elegans* genome and also mammalian databases for homologous proteins, revealing two other *C. elegans* homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRAC-Subsequent BLAST searches picked up mouse EST like currents were best defined. sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures

5

10

15

20

25

30

Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promers were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

5

10

15

20

25

30

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y., N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in <u>Capacitative Calcium Entry</u>, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2+ indicators with greatly improved fluorescence properties, J. Biol

-45-

5

10

15

20

25

30

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca2⁺ indicators with greatly improved fluorescence properties, <u>J Biol Chem 260</u>, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging [Ca2+]i in single cells, <u>Prog Clin Biol Res 210</u>, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., *EMBO J*, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (*Cell*, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure 1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

5

10

15

20

25

30

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

-47-

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in Xenopus oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: Pharmacologic Behavior of SOC/CRAC

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: Gene targeting

5

10

15

20

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectivelly refered to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:

SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

-48-

Claims

1. An isolated nucleic acid molecule, comprising:

5

10

15

20

- (a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;
- (b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;
- (c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and
 - (d) complements of (a), (b) or (c).
- 2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.
- 3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.
- 4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.
- 5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.
- 6. An isolated nucleic acid molecule selected from the group consisting of
- (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,
- (b) complements of (a),
 provided that the unique fragment includes a sequence of contiguous nucleotides which is not
 identical to any sequence selected from a sequence group consisting of
 - (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,
 - (2) complements of (1), and
 - (3) fragments of (1) and (2).

5.

WO 00/40614 PCT/US99/29996

- 7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:
 - (1) at least two contiguous nucleotides nonidentical to the sequence group,
 - (2) at least three contiguous nucleotides nonidentical to the sequence group,
 - (3) at least four contiguous nucleotides nonidentical to the sequence group,
 - (4) at least five contiguous nucleotides nonidentical to the sequence group,
 - (5) at least six contiguous nucleotides nonidentical to the sequence group,
 - (6) at least seven contiguous nucleotides nonidentical to the sequence group.
- 8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.
- 9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.
 - 10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.
 - 11. A host cell transformed or transfected with the expression vector of claim 10.
- 20 12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.
 - 13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.
- 25 14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

- 15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3. 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.
- 16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.
- 5 17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.
 - 18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.
 - 19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.
 - 20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:
 - a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;
 - b) detecting the presence of the complex;

10

15

20

- c) isolating the SOC/CRAC molecule from the complex; and
- d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.
- 21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.
- 22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

- WO 00/40614 PCT/US99/29996 -51-
- 23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3. SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.
- 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

5

10

15

20

- a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;
- b) detecting a Ca²⁺ concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.
- 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:
- a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and
- b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.
- 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.
- 27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.
- 28. The method of claim 25, wherein the test sample is tissue.
- 29. The method of claim 25, wherein the test sample is a biological fluid.

WO 00/40614 PCT/US99/29996 -52-

- 30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).
- 31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to SOC/CRAC and polyclonal antisera to SOC/CRAC.
- 32. A kit, comprising a package containing:

5

10

15

20

25

- an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and
- a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.
- 33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.
- 34. A pharmaceutical composition comprising:
- a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and
 - a pharmaceutically acceptable carrier.
- 35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.
- 36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:
- a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;
- b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and
- c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC kinase activity.
- 37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

1/1

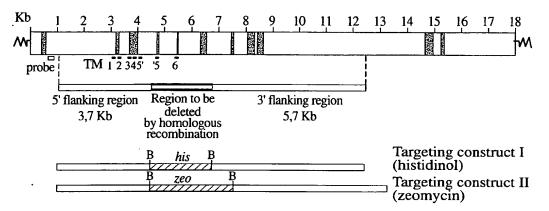


Fig. 1

-1-

```
SEQUENCE LISTING
      <110> Beth Israel Deaconess Medical Center, Inc.
            Scharenberg, Andrew
      <120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY
      <130> B0662/7026WO/ERP/KA
      <150> U.S. 60/114,220
      <151> 1998-12-30
      <150> U.S. 60/120,018
      <151> 1999-01-29
      <150> U.S. 60/140,415
      <151> 1999-06-22
      <160> 32
      <170> FastSEQ for Windows Version 3.0
      <210> 1
      <211> 1212
      <212> DNA
      <213> Homo Sapiens
      <400> 1
                                                                        60
gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaaag ccacttggaa
                                                                       120
actggaacca aagatcaaga aactgtttgc tctaaagcta cagaaggaga taatacagaa
tttggagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca
                                                                       180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt
                                                                       240
                                                                       300
tctcttgctg gatttactga ctgtcacaga acttccattc ctgttcattc aaaacaagaa
                                                                       360
aaaatcagta gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta
ataccggttt gtagatttca actaaacaga tatatattat taaatacatt aaactttttt
                                                                       420
agataagatc tacaaagtgg tgatatttgg gactatatca aaaattcaaa aaaatttttc
                                                                       540
ttaagaaaac tgactttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg
aagctaatta tatgaactca tttccagcta actgcgataa tgattgattt tgcaaattcc
                                                                       660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt
                                                                       780
acattcatat atatctgttc tggcatagag agattatgtt gttaaaatca tgttattgaa
                                                                       840
                                                                       900
aaaagttatt tcagtgggga aagaggttag ttaacaaaga gattcacagt aacaaatcct
cetttetgga gggaetette etgaecetga getgeacaae tttgeaacaa attaaageet
                                                                       960
aaccqaaqat qacctcacaa tqqcaattta qaactcatgq qagtcaactt acataaacgq
                                                                      1020
                                                                      1080
tatttgattt ctgataagat agtggaatta ttggttatag atgacaaaat aagtatgttt
aaagtgatga tggacataaa aaagttttaa atataaaaca tgagaaaaga aggagatact
                                                                      1140
                                                                      1200
attcaaaaag actggcaaat ttgaaaaact agaaataaaa aaaaaaaaa aaaatgagcg
                                                                      1212
gccgcaagct tt
     <210> 2
     <211> 141
     <212> PRT
     <213> Homo Sapiens
```

Ala Arg Gly Lys Phe Phe Val Ser Thr Pro Ser Gln Pro Ser Cys Lys

Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys

25

10

15

30

<400> 2

```
Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe Val Gly His Arg
                         40
                                                45
Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Lys
                        55
                                            60
    50
Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu Lys Arg Val Ser
                    70
                                        75
Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser Ile Pro Val His
                                    90
Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro Ser Thr Glu Asp Thr His
                             105
          100
Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Val Cys Arg Phe Gln Leu
                            120
                                                125
        115
Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn Phe Phe Arg
                        135
    130
      <210> 3
      <211> 739
      <212> DNA
      <213> Homo Sapiens
      <220>
      <221> unsure
      <222> (5)...(5)
      <223> UNKNOWN
      <221> unsure
      <222> (21)...(22)
      <223> UNKNOWN
      <221> unsure
      <222> (29)...(29)
      <223> UNKNOWN
      <400> 3
                                                                        60
tegantaggg gtettecace nneatactng gatgatggtt ggtgaagtet atgeatacga
aattgatgtg tgtgcaaacg attctgttat ccctcaaatc tgtggtcctg ggacgtggtt
                                                                       120
qactccattt cttcaagcag tctacctctt tgwacagtat atcattatgg ttaatcttct
                                                                       180
tattgcattt ytcaacaatg tgtatttaca agtgaaggca atttccaata ttgyatggaa
                                                                       240
                                                                       300
gtaccagcgt tatcatttta ttatggctta tcatgagaaa ccagttctgc ctcctccact
                                                                       360
tatcattctt agccatatag tttctctgtt ttgctgcata tgtaagagaa gaaagaaaga
taaqacttcc qatqqaccaa aacttttctt aacaqaagaa gatcaaaaga aacttcatga
                                                                       420
ttttgaagag cagtgtgttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg
                                                                       480
gagtgaagag agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa
                                                                       540
                                                                       600
agaagttgga gatccgtgtc aactacataa aaagatcatt acaatcatta gattctcaaa
                                                                       660
ttggccattt gcaagatctt tcagccctga cggtagatac attaaaaaca ctcactggcc
                                                                       720
aaaagcgtcg gaagctagca aagttcataa tgaaatcaca cgagaactga gcatttccaa
                                                                       739
acacttggct caaaacctt
      <210> 4
      <211> 235
      <212> PRT
      <213> Homo Sapiens
      <220>
      <221> UNSURE
      <222> (41)...(41)
      <223> UNKNOWN
      <221> UNSURE
      <222> (54)...(54)
```

<221> unsure <222> (375)...(375)

<223> g or c

PCT/US99/29996 WO 00/40614 -3-

```
<223> UNKNOWN
     <221> UNSURE
     <222> (68)...(68)
     <223> UNKNOWN
     <400> 4
Met Met Val Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys Ala Asn
Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu Thr Pro
                               25
       20
Phe Leu Gln Ala Val Tyr Leu Phe Xaa Gln Tyr Ile Ile Met Val Asn
                        40
Leu Leu Ile Ala Phe Xaa Asn Asn Val Tyr Leu Gln Val Lys Ala Ile
                   55
Ser Asn Ile Xaa Trp Lys Tyr Gln Arg Tyr His Phe Ile Met Ala Tyr
65 70 75 80
                  70
                                     75
His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile
85 90 95
              85
Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp Lys Thr
                              105
          100
Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu
                       120
                                              125
      115
His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu Lys Asp
                                         140
                     135
 130
Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr Phe Glu
                                    155
                  150
Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp Pro Cys 165 170 175
Gln Leu His Lys Lys Ile Ile Thr Ile Ile Arg Phe Ser Asn Trp Pro
                              185 . 190
          180
Phe Ala Arg Ser Phe Ser Pro Asp Gly Arg Tyr Ile Lys Asn Thr His
195 200 205
Trp Pro Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg
210 215 220
                   215
Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn
                   230
     <210> 5
     <211> 1579
      <212> DNA
      <213> Homo Sapiens
     <220>
     <221> unsure
      <222> (368)...(368)
     <223> g or c
     <221> unsure
      <222> (372)...(372)
     <223> g or c
     <221> unsure
     <222> (374)...(374)
     <223> g or a
```

PCT/US99/29996 WO 00/40614

60

-4-<221> unsure <222> (387)...(387) <221> unsure <222> (482)...(482) <400> 5 acgtcgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc ggcatggtgt 120 tgtaaataca cttagctcct ctcttcctca aggtgatctt gaaagtaata atccttttca 180 tigtaatatt ttaatgaaag atgacaaaga tccccagtgt aatatatttg gtcaagactt acctgcagta ccccagagaa aagaatttaa ttttccagag gctggttcct cttctggtgc 240 300 cttattccca agtgctqttt cccctccaga actgcgacag agactacatg gggtagaact cttaaaaata tttaataaaa atcaaaaatt aggcagttca tctactagca taccacatct 360 420 qtcatccsca csarscaaat tttttqntag tacaccatct cagccaagtt gcaaaagcca 480 cttggaaact ggaaccaaag atcaagaaac tgtttgctct aaagctacag aaggagataa 540 tncagaattt ggagcatttg taggacacag agatagcatg gatttacaga ggtttaaaga aacatcaaac aagataaaaa tactatccaa taacaatact totgaaaaca otttgaaacg 600 agtgagttct cttgctggat ttactgactg tcacagaact tccattcctg ttcattcaaa 660 acaagaaaaa atcagtagaa ggccatctac cgaagacact catgaagtag attccaaagc 720 agctttaata ccggtttgta gatttcaact aaacagatat atattattaa atacattaaa 780 cttttttaga taagatctac aaagtggtga tatttgggac tatatcaaaa attcaaaaaa 840 900 atttttctta agaaaactga ctttagcata gtagcagtta cagaaaagtt tcttacagtg aatagtcagg aattttaaag aaaaatttat gcagaataaa ggcaggaatc tctttttgtt 960 tgaattgaag ctaattatat gaactcattt ccagctaact gcgataatga ttgattttgc 1020 1080 aaattccctt taaaagcaca cactgacaag acaaaaagct caggaaaaagg cagaaaaatt 1140 actcctttat aatcaagtat tatatataag tcagtgctca taattttgct caagaaaata ttgacttaca ttcatatata tctgttctgg catagagaga ttatgttgtt aaaatcatgt 1200 tattgaaaaa agttatttca gtggggaaag aggttagtta acaaagagat tcacagtaac 1260 aaatcctcct ttctggaggg actcttcctg accctgagct gcacaacttt gcaacaaatt 1320 aaagcctaac cgaagatgac ctcacaatgg caatttagaa ctcatgggag tcaacttaca 1380 1440 taaacggtat ttgatttctg ataagatagt ggaattattg gttatagatg acaaaataag tatgtttaaa gtgatgatgg acataaaaaa gttttaaata taaaacatga gaaaagaagg 1500 1560 agatactatt caaaaagact ggcaaatttg aaaaactaga aataaaaaaa aaaaaaaaa 1579 atgageggee geaagettt <210> 6 <211> 243 <212> PRT <213> Homo Sapiens <220> <221> UNSURE <222> (103)...(105) <223> UNKNOWN <221> UNSURE <222> (109)...(109) <223> UNKNOWN <221> UNSURE <222> (141)...(141) <223> UNKNOWN <400> 6 Val Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn 5 10 Asn Pro Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln 20 25

Cys Asn Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu 40

```
-5-
Phe Asn Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser
                     55
Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu
                 70 75
Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser
              85
                                  90
Ile Pro His Leu Ser Ser Xaa Xaa Xaa Lys Phe Phe Xaa Ser Thr Pro
           100
                              105
                                                  110
Ser Gln Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln
      115
                          120
                                     125
Glu Thr Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Xaa Glu Phe Gly
                       135
                                           140
Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu
                  150
                                      155
Thr Ser Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn
                                   170
                                                      175
               165
Thr Leu Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg
                            185
Thr Ser Ile Pro Val His Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro
                           200
       195
Ser Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro
                    215
Val Cys Arg Phe Gln Leu Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn
Phe Phe Arg
      <210> 7
      <211> 3532
      <212> DNA
      <213> Mus Musculus
      <220>
      <221> unsure
      <222> (2420)...(2420)
      <223> unknown
     <221> unsure
      <222> (2434)...(2434)
      <223> unknown
     <221> unsure
      <222> (2461)...(2461)
      <223> unknown
     <221> unsure
     <222> (2466)...(2466)
     <223> unknown
     <221> unsure
     <222> (2470)...(2470)
     <223> unknown
     <400> 7
attatggctt atcatgaaaa accagtcctg cctcctcctc ttatcatcct cagccatata
gtttcactgt tttgctgtgt atgcaaaaga agaaagaaag ataagacttc cgatgggcca
                                                                    120
                                                                    180
aaacttttct taacagaaga agatcaaaag aaactccatg attttgaaga gcagtgtgtt
gagatgtact ttgatgagaa agatgacaaa ttcaattctg ggagtgaaga gagaatccgg
                                                                    240
gtcacttttg aaagagtgga gcagatgagc attcagatta aagaagttgg agatcgtgtc
                                                                    300
```

aactacataa aaagatcatt acagtettta gatteteaaa ttggteatet geaagatete

PCT/US99/29996 WO 00/40614 -6-

tcagccctaa cagtagatac attgaaaaca cttacagccc agaaagcttc agaagctagt 420 480 aaagtgcaca atgagatcac acgagaattg agtatttcca aacacttggc tcagaatctt 540 attgatgatg ttcctgtaag acctttgtgg gaagaaccta gtgctgtaaa cacactgagt tectetete etcaaggtga tegggaaagt aataateett ttetttgtaa tattttatg 600 660 aaagatgaaa aagacccca atataatctg tttggacaag atttgcccgt gataccccag agaaaagaat tcaacattcc agaggctggt tcctcctgtg gtgccttatt cccaagtgct 720 gtttctcccc cagaattacg acagagacga catggggtag aaatgttaaa aatatttaat 780 aaaaatcaaa aattaggcag ttcacctaat agttcaccac atatgtcctc cccaccaacc 840 aaattttctg tgagtacccc atcccagcca agttgcaaaa gtcacttgga atccacaacc 900 960 aaagatcaag aacccatttt ctataaagct gcagaagggg ataacataga atttggagca tttgtgggac acagagatag tatggactta cagaggttta aagaaacatc aaacaaaata 1020 1080 agagaactgt tatctaatga tactcctgaa aacactctga aacatgtggg tgctgctgga tataqtqaat qttqtaaqac ttctacttct cttcactcgg tqcaaqcaga aaqctqtaqt 1140 agaagagcgt cgacggaaga ctctccagaa gtcgattcta aagcagcttt gttaccggat 1200 tggttacgag atagaccatc aaacagagaa atgccatctg aaggaggaac attaaatggt 1260 1320 cttgcttctc catttaagcc cgttttggat acaaattact attattcagc tgtggaaaga aataacctga tgaggttgtc acagagtatt cccttcgttc ctgtacctcc acgaggcgag 1380 1440 cctgtcacag tgtaccgtct ggaggagagt tctcccagta tactgaataa cagcatgtct tcatggtctc agctaggcct ctgtgccaaa attgagtttt taagtaaaga ggaaatggaa 1500 ggtggtttac gaagagcagt caaagtgctg tgtacctggt cagagcacga tatcctgaag 1560 teagggeate tetatateat taagteattt etteetgagg tgataaacae atggteaage 1620 1680 atttataaag aagatacggt tctacatctc tgtctcagag aaatacaaca acagagagca 1740 gcacaaaagc tcacatttgc ctttaatcag atgaaaccca aatccatacc atattctcca aggttccttg aagttttcct gttgtactgc cattcagcag ggcagtggtt tgctgtagaa 1800 gagtgcatga ctggtgaatt tagaaaatac aacaacaata atggtgatga aatcattcct 1860 1920 acaaatactc tagaagagat catgctagcc tttagccact ggacctatga atataccaga ggggagttac tggtacttga cttacaagga gtgggagaaa acttgactga cccatctgta 1980 ataaaagctg aagaaaaaag atcctgtgac atggtttttg gccctgccaa tctaggagaa 2040 gatgcaataa aaaacttcaa gagccaaaca tccactgtaa ttcttgctgt cgaaaqctta 2100 aacttcccag atttgaagag gaatgactac acgcccttga taaaattata tttcctcagg 2160 atgagtcatc agatttgaat cttcaatctg gaaattccac caaagaatca gaagcaacaa 2220 attotgttcg totgatgtta tagtgctgag toattggttt ttgcctacac ttcacaaaag 2280 tgtaactgtc agttttcctt tcgggggaat tgatgatata ggaagatgtg tgcaaaatga 2340 gcttqctqqc cccacacata qtctaqaqqt aatqttctca ttgaaaaacg cctqqaggtg 2400 gaggetgeag atgccagtgn aaagtgctag etgncagaga gtcagtgete tegggetggt 2460 naaggneggn accettgetg etgagagtgg tggttetett cacetggtge aggaceatta accaaagtea agtetteaga tttgattgge tgeteagtea eageecatte agetaaggaa 2520 2580 actaeattgc gcagcttttt aeatggctga agtcttcctc agtttgtgct ctatgataat 2640 2700 gatgttagct ctcaactagg tgtttgtggc cacgggagaa ctactcctta caattttgct tcacaggcat gttacaaagc ctgcactgaa aaccgtttgt cttccctctc tccctccctc 2760 2820 ttttccctgt agtattgagg atcaaaccca gggcctcatg aagaccattt tctaagagac attttattta agaatcaact atagagtcta tgtttatgga tacagccagt ttttgttaaa 2880 caaaacctga attgtgcaaa agggtttttt aacatttatc aatgttaagt aaaagaaagc 2940 catgataaat aagaattaac tcactgttca atgggtgttt cctgtgagga aggttacagt 3000 tgtaacagcc tgcagttgca tacatctcca aagatttaca gacttagtgt atcaaatcag 3060 agtgtcatgt gagctctcac attgaaaatt ctataggaat gtgtcaatgt gaattctatt 3120 tctggtactt aagaaatcag ttgttggatt atccttatac agtataggga gatcacaata 3180 caactttatg ccaataaaat ctaacttaat tgcccagata tttttgcata tttagcaaca 3240 agaaaagctt atcatttgac tcaagtttta tgctttctct ttctttcat ttcctaggta 3300 ctaattttaa tttttatttg gaaggagcag tgtaaagctt acttgtattc aatagtgtat 3360 3420 ctcatagata cagacaaggc cgcagagata agctgttaaa tagtgtttaa tgttgatgtg 3480 gagagaaagg tgtattactt aaaaatacta taccatatac gttttgtata tcattaaatc 3532

<210> 8

<211> 475

<212> PRT

<213> Mus Musculus

<400> 8

Ile Met Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile Val Ser Leu Phe Cys Cys Val Cys Lys Arg Arg Lys Lys Asp Lys Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asp Glu Lys Asp Asp Lys Phe Asn Ser Gly Ser Glu Glu Arg Ile Arg 65 70 75 80 Val Thr Phe Glu Arg Val Glu Gln Met Ser Ile Gln Ile Lys Glu Val Gly Asp Arg Val Asn Tyr Ile Lys Arg Ser Leu Gln Ser Leu Asp Ser Gln Ile Gly His Leu Gln Asp Leu Ser Ala Leu Thr Val Asp Thr Leu Lys Thr Leu Thr Ala Gln Lys Ala Ser Glu Ala Ser Lys Val His Asn Glu Ile Thr Arg Glu Leu Ser Ile Ser Lys His Leu Ala Gln Asn Leu Ile Asp Asp Val Pro Val Arg Pro Leu Trp Glu Glu Pro Ser Ala Val Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Arg Glu Ser Asn Asn Pro Phe Leu Cys Asn Ile Phe Met Lys Asp Glu Lys Asp Pro Gln Tyr Asn Leu Phe Gly Gln Asp Leu Pro Val Ile Pro Gln Arg Lys Glu Phe 210 215 220 Asn Ile Pro Glu Ala Gly Ser Ser Cys Gly Ala Leu Phe Pro Ser Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Arg His Gly Val Glu Met Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Pro Asn Ser Ser Pro His Met Ser Ser Pro Pro Thr Lys Phe Ser Val Ser Thr Pro Ser Gln Pro Ser Cys Lys Ser His Leu Glu Ser Thr Thr Lys Asp Gln Glu Pro Ile Phe Tyr Lys Ala Ala Glu Gly Asp Asn Ile Glu Phe Gly Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser Asn Lys Ile Arg Glu Leu Leu Ser Asn Asp Thr Pro Glu Asn Thr 340 345 Leu Lys His Val Gly Ala Ala Gly Tyr Ser Glu Cys Cys Lys Thr Ser Thr Ser Leu His Ser Val Gln Ala Glu Ser Cys Ser Arg Arg Ala Ser Thr Glu Asp Ser Pro Glu Val Asp Ser Lys Ala Ala Leu Leu Pro Asp Trp Leu Arg Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Gly Gly Thr Leu Asn Gly Leu Ala Ser Pro Phe Lys Pro Val Leu Asp Thr Asn Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln Ser Ile Pro Phe Val Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val 455 460 Tyr Pro Ser Gly Gly Arg Val Leu Pro Val Tyr

```
<210> 9
<211> 5433
<212> DNA
<213> Mus Musculus
<220>
<221> unsure
<222> (5094)...(5094)
<223> unknown
```

<400> 9

ggctgaaaga gcctgagctg tgcctctcca ttccactgct gtggcagggt cagaaatctt 60 120 ggatagagaa aaccttttgc aaacgggaat gtatctttgt aattcctagc acgaaagact ctaacaggtg ttgctgtggc cagttcacca accagcatat cccccctctg ccaagtgcaa 180 240 cacccagcaa aaatgaagag gaaagcaaac aggtggagac tcagcctgag aaatggtctg ttgccaagca cacccagagc tacccaacag attcctatgg agttcttgaa ttccagggtg 300 360 gcggatattc caataaagcc atgtatatcc gtgtatccta tgacaccaag ccagactcac tgctccatct catggtgaaa gattggcagc tggaactccc caagetetta atatctgtgc 420 atggaggcct ccagaacttt gagatgcagc ccaagctgaa acaagtcttt gggaaaggcc 480 tgatcaagge tgetatgace accggggeet ggatetteae egggggtgte ageacaggtg 540 600 ttatcagcca cgtaggggat gccttgaaag accactcctc caagtccaga ggccgggttt gtgctatagg aattgctcca tggggcatcg tggagaataa ggaagacctg gttggaaagg 660 720 atqtaacaaq aqtqtaccaq accatgtcca accetctaag taagetetet gtgeteaaca 780 acteceacae ceaetteate etggetgaca atggeaceet gggeaagtat ggegeegagg tgaagetgeg aaggetgetg gaaaageaca teteeeteea gaagateaac acaagaetgg 840 900 ggcagggcgt gcccctcgtg ggtctcgtgg tggagggggg ccctaacgtg gtgtccatcg tottggaata cotgoaagaa gagootooca toootgtggt gatttgtgat ggcagoggac 960 1020 gtgcctcgga catcctgtcc tttgcgcaca agtactgtga agaaggcgga ataataaatg agtccctcag ggagcagctt ctagttacca ttcagaaaac atttaattat aataaggcac 1080 1140 aatcacatca gctgtttgca attataatgg agtgcatgaa gaagaaagaa ctcgtcactg 1200 tgttcagaat gggttctgag ggccagcagg acatcgagat ggcaatttta actgccctgc 1260 tgaaaggaac aaacgtatct gctccagatc agctgagctt ggcactggct tggaaccgcg 1320 tqqacataqc acqaaqccaq atctttqtct ttgggcccca ctggacgccc ctgggaagcc 1380 tggcacccc gacggacage aaagccacgg agaaggagaa gaagccaccc atggccacca ccaagggagg aagaggaaaa gggaaaggca agaagaaagg gaaagtgaaa gaggaagtgg 1440 1500 aggaagaaac tgaccccgg aagatagagc tgctgaactg ggtgaatgct ttggagcaag cgatgctaga tgctttagtc ttagatcgtg tcgactttgt gaagctcctg attgaaaacg 1560 1620 gagtgaacat gcaacacttt ctgaccattc cgaggctgga ggagctctat aacacaagac tgggtccacc aaacacactt catctgctgg tgagggatgt gaaaaagagc aaccttccgc 1680 1740 ctgattacca catcagcctc atagacatcg ggctcgtgct ggagtacctc atgggaggag cctaccqctq caactacact cqqaaaaact ttcqqaccct ttacaacaac ttqtttqgac 1800 1860 caaagaggcc taaagctctt aaacttctgg gaatggaaga tgatgagcct ccagctaaag ggaagaaaaa aaaaaaaag aaaaaggagg aagagatcga cattgatgtg gacgaccctg 1920 1980 ccgtgagtcg gttccagtat cccttccacg agctgatggt gtgggcagtg ctgatgaaac gccagaaaat ggcagtgttc ctctggcagc gaggggaaga gagcatggcc aaggccctgg 2040 2100 tggcctgcaa gctctacaag gccatggccc acgagtcctc cgagagtgat ctggtggatg acatetecea ggaettggat aacaatteea aagaettegg eeagettget ttggagttat 2160 2220 tagaccagte etataageat gacgageaga tegetatgaa acteetgace tacgagetga 2280 aaaactqqaq caactcqacc tqcctcaaac tggccgtggc agccaaacac cgggacttca 2340 ttgctcacac ctgcagccag atgctgctga ccgatatgtg gatgggaaga ctgcggatgc ggaagaaccc cggcctgaag gttatcatgg ggattcttct acccccacc atcttgtttt 2400 tggaatttcg cacatatgat gatttctcgt atcaaacatc caaggaaaac gaggatggca 2460 aagaaaaaga agaggaaaat acggatgcaa atgcagatgc tggctcaaga aagggggatg 2520 aggagaacga qcataaaaaa cagagaagta ttcccatcgg aacaaagatc tgtgaattct 2580 2640 ataacgcgcc cattgtcaag ttctggtttt acacaatatc atacttgggc tacctgctgc tgtttaacta cgtcatcctg gtgcggatgg atggctggcc gtccctccag gagtggatcg 2700 2760 tcatctccta catcgtgagc ctggcgttag agaagatacg agagatcctc atgtcagaac caggcaaact cagccagaaa atcaaagttt ggcttcagga gtactggaac atcacagatc 2820 2880 tegtggeeat trecacatte atgattggag caattetteg cetacagaac cageeetaca

tgggctatgg ccgggtgatc tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg 2940 3000 acatetttgg tgtcaacaag tatetgggge cataegtgat gatgattgga aagatgatga tcgacatgct gtactttgtg gtcatcatgc tggtcgtgct catgagtttc ggagtagccc 3060 qtcaaqccat totqcatoca qaggagaago cotottggaa actggcccga aacatottot 3120 3180 acatgeecta etggatgate tatggagagg tgtttgeaga ecagatagae etetaegeea 3240 tqqaaattaa teeteettgt ggtgagaace tatatgatga ggagggcaag eggetteete 3300 cctgtatccc cggcgcctgg ctcactccag cactcatggc gtgctatcta ctggtcgcca acatectget ggtgaacetg etgattgetg tgttcaacaa taetttettt gaagtaaaat 3360 3420 caatatccaa ccaggtgtgg aagttccagc gatatcagct gattatgaca tttcatgaca ggccagtcct gccccaccg atgatcattt taagccacat ctacatcatc attatgcgtc 3480 3540 tcagcggccg ctgcaggaaa aagagagaag gggaccaaga ggaacgggat cgtggattga agctcttcct tagcgacgag gagctaaaga ggctgcatga gttcgaggag cagtgcgtgc 3600 3660 aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcatccggg tcacttctga aagagttgaa aatatgtcaa tgaggttgga agaaatcaat gaaagagaaa 3720 3780 cttttatgaa aacttccctg cagactgttg accttcgact tgctcagcta gaagaattat ctaacagaat ggtgaatgct cttgaaaatc ttgcgggaat cgacaggtct gacctgatcc 3840 aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccgg caaagcagca 3900 tcaatagcgc tgatggctac agcttgtatc gatatcattt taacggagaa gagttattat 3960 4020 ttgaggatac atctctctc acgtcaccag ggacaggagt caggaaaaaa acctgttcct 4080 tccgtataaa ggaagagaag gacgtgaaaa cgcacctagt cccagaatgt cagaacagtc ttcacctttc actgggcaca agcacatcag caaccccaga tggcagtcac cttgcagtag 4140 4200 atgacttaaa gaacgctgaa gagtcaaaat taggtccaga tattgggatt tcaaaggaag atgatgaaag acagacagac tctaaaaaaag aagaaactat ttccccaagt ttaaataaaa 4260 4320 cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa cgacaaatat agaaggcact atttcctatc ccctggaaga aaccaaaatt acacgctatt 4380 4440 teccegatga aacgateaat gettgtaaaa caatgaagte cagaagette gtetatteee 4500 qqqqaaqaaa gctggtcggt ggggttaacc aggatgtaga gtacagttca atcacggacc 4560 agcaattgac gacggaatgg caatgccaag ttcaaaagat cacgcgctct catagcacag atatteetta cattgtgteg gaagetgeag tgeaagetga geaaaaagag cagtttgeag 4620 atatgcaaga tgaacaccat gtcgctgaag caattcctcg aatccctcgc ttgtccctaa 4680 ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg 4740 4800 gatteceate teteaggtea aaaagtttae atggacatee taggaatgtg aaateeatte agggaaagtt agacagatct ggacatgcca gtagtgtaag cagcttagta attgtgtctg 4860 gaatgacagc agaagaaaaa aaggttaaga aagagaaagc ttccacagaa actgaatgct 4920 agtetgtttt gtttetttaa tttttttttt taacagteag aaacceacta atgggtgtea 4980 5040 tettggeeca teetaaacae atmteeaatt teetaaaaae atttteeett aaaaaatttt ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatatgt 5100 5160 tagtaggett agttttttea gttgeagtag tateaaatga aagtgatgat aetgtaaega agataaattg gctaatcagt atacaagatt atacaatctc tttattactg agggccacca 5220 aatageetag gaagtgeeet egageaetga agteaceatt aggteaetea agaagtaage 5280 aactagctgg gcacagtggc tcatgcctgt aatcctagca ctttgggagg ccaaggcaga 5340 aagatagett gagteeagga gtttgagaee ageetgggea acatagtgat acceeatete 5400 5433 ttaaaaaaa aaaaaaaaa ctgccctcgt gcc

<210> 10 <211> 1533 <212> PRT

<213> Mus Musculus

-10-

								•	-10-						
Ile 65	Phe	Thr	Gly	Gly	Val 70	Ser	Thr	Gly	Val	Ile 75	Ser	His	Val	Gly	Asp 80
	Leu	Lys	Asp	His 85	Ser	Ser	Lys	Ser	Arg 90	Gly	Arg	Val	Cys	Ala 95	Ile
Gly	Ile	Ala	Pro 100	Trp	Gly	Ile	Val	Glu 105	Asn	Lys	Glu	Asp	Leu 110	Val	Gly
Lys	Asp	Val 115	Thr	Arg	Val	Tyr	Gln 120	Thr	Met	Ser	Asn	Pro 125	Leu	Ser	Lys
Leu	Ser 130	Val	Leu	Asn	Asn	Ser 135	His	Thr	His	Phe	Ile 140	Leu	Ala	Asp	Asn
145					150					155				Leu	160
				165					170					Gln 175	
			180					185					190	Val	
		195					200					205		Val	
	210					215					220			His	
225					230					235				Gln	240
				245					250					Ser 255	
			260					265					270	Leu	
		275					280					285		Met	
	290					295					300			Asp	
305					310					315				Ser	320
				325					330					Ala 335	
			340					345					350	Met Gly	
		355					360					365		Glu	
	370					375					380			Leu	
385					390					395				Val	400
				405					410					415 Asn	
			420					425					430	Val	
		435					440					445			
	450					455					460				Gly
465					470					475				Tyr	480
				485					490					Lys 495	
			500					505					510	Pro	
		515					520					525		Asp	
Asp	Val 530	Asp	Asp	Pro	Ala	Val 535	Ser	Arg	Phe	Gln	Tyr 540	Pro	rne	His	GIU

-||-

545			Trp		550					555					560
	_		Arg	565					570					575	
_		_	Lys 580					585					590		
Asp	Asp	Ile 595	Ser	Gln	Asp	Leu	Asp 600	Asn	Asn	Ser	Lys	Asp 605	Phe	Gly	Gln
Leu	Ala 610	Leu	Glu	Leu	Leu	Asp 615	Gln	Ser	Tyr	Lys	His 620	Asp	Glu	Gln	Ile
Ala 625	Met	Lys	Leu	Leu	Thr 630	Tyr	Glu	Leu	Lys	Asn 635	Trp	Ser	Asn	Ser	Thr 640
-		•	Leu	645				_	650					655	
	_		Gln 660					665					670		
		675	Asn				680					685			
	690		Leu			695					700				
705			Lys		710					715					720
	_		Asn	725					730					735	
			Lys 740					745					750		
		755	Ala				760					765			
	770		Leu			775					780				
785	-		Ser		790		_			795					800
			Glu	805					810					815	
			Lys 820					825					830		
		835	Ala				840					845			
	850	•	Pro			855					860				
865			Trp		870					875					880
			Pro	885					890					895	
			Val 900					905					910		
	_	915	Ala				920					925			
	930		Ile			935					940				
945			Gln		950					955					960
			Leu	965					970					975	
			Trp 980					985					990		
		995	Leu				1000)				1005	5		
Phe	Phe 1010		Val	Lys	Ser	Ile 1015		Asn	Gln	Val	Trp 1020		Phe	Gln	Arg

-12-

Tyr 1025															
		Leu	Ile	Met			His	Asp	Arg			Leu	Pro	Pro	Pro 104
			_	_	1030		_			1035		_	_	_	
			Leu	1045	5				1050)				1055	5
Ara	Cvs	Ara	Lys	Lys	Arq	Glu	Gly	Asp	Gln	Glu	Glu	Arg	Asp	Arg	Gly
	- 4	_	1060		_		-	1069				_	1070		_
Leu	Lys		Phe		Ser	Asp		Glu		Lys	Arg		His	-	Phe
		1075					1080					1085			
Glu	Glu 1090		Суѕ	Val	Gln	Glu 1095		Phe	Arg	Glu	Lys 1100		Asp	Glu	Gln
Gln			Ser	Asp	Glu	Ara	He	Ara	Val	Thr	Ser	Glo	Ara	Val	Glu
1105		001		1.02	1110			9		1115			5		112
		C	Met	7-~			Clu	T10	700			Glu	Th ~	Dha	
			•	1125	5				1130)	-			1135	5
Lys	Thr	Ser	Leu	Gln	Thr	Val	Asp	Leu	Arg	Leu	Ala	Gln	Leu	Glu	Glu
			1140					1145					1150		
ī.e.u	Ser	Asn	Arg	Met	Val	Asn	Ala	Len	Glu	Asn	Leu	Ala	Glv	Ile	Asp
200	001	1155					1160					1165			
	.			T 1 -	C1-	7.1			71	71.	C			C	C1
Arg		_	Leu	тте	GIn			ser	Arg	AIA			GIU	Cys	GIU
	1170					1175					1180				
Ala	Thr	Tyr	Leu	Leu	Arg	Gln	Ser	Ser	Ile	Asn	Ser	Ala	Asp	Gly	Tyr
1185	5				1190)				1195	5				120
Ser	Leu	Tyr	Arg	Tyr	His	Phe	Asn	Gly	Glu	Glu	Leu	Leu	Phe	Glu	Asp
		•	-	1205				_	1210					1215	
Thr	Ser	T.e.n	Ser			Pro	Glv	Thr			Ara	Lvs	Lvs	Thr	Cvs
*,***	DCI	пси	1220		001		- 1	1225			9	_,	1230	`	0,70
~	Db -	70			C1	<i>~</i> 1	T		-	T	m 10 m	11: -	-		Dro
Ser	Pne		Ile	гàг	GIU	GIU			vaı	ьys	Thr			vaı	PIO
		1235				_	1240					1245		_	
Glu	Cys	Gln	Asn	Ser	Leu			Ser	Leu	Gly			Thr	Ser	Ala
	1250					1255					1260				
Thr	Pro	Asp	Gly	Ser	His	Leu	Ala	Val.	Asp	Asp	Leu	Lys	Asn	Ala	Glu
1265					1270					1275					128
		Lvs	Leu	Gly			Ile	Gly	Ile			Glu	Asp	Asp	
		Lys	Leu		Pro		Ile	Gly	Ile 1290	Ser		Glu	Asp	Asp 1295	Glu
Glu	Ser			1285	Pro	Asp			1290	Ser	Lys			1295	Glu 5
Glu	Ser		Asp	1285 Ser	Pro	Asp		Glu	1290 Thr	Ser	Lys		Ser	1295 Leu	Glu 5
Glu Arg	Ser Gln	Thr	Asp 1300	1285 Ser)	Pro Lys	Asp Lys	Glu	Glu 1305	1290 Thr	Ser) Ile	Lys Ser	Pro	Ser 1310	1295 Leu)	Glu S Asn
Glu Arg	Ser Gln	Thr Asp	Asp 1300 Val	1285 Ser)	Pro Lys	Asp Lys	Glu Gln	Glu 1305 Asp	1290 Thr	Ser) Ile	Lys Ser	Pro Val	Ser 1310 Gln	1295 Leu)	Glu S Asn
Glu Arg Lys	Ser Gln Thr	Thr Asp 1315	Asp 1300 Val	1285 Ser) Ile	Pro Lys His	Asp Lys Gly	Glu Gln 1320	Glu 1305 Asp	1290 Thr b Lys	Ser) Ile Ser	Lys Ser Asp	Pro Val 1325	Ser 1310 Gln	1295 Leu) Asn	Glu Asn Thr
Glu Arg Lys	Ser Gln Thr Leu	Thr Asp 1315 Thr	Asp 1300 Val	1285 Ser) Ile	Pro Lys His	Asp Lys Gly Thr	Glu Gln 1320 Asn	Glu 1305 Asp	1290 Thr b Lys	Ser) Ile Ser	Lys Ser Asp Thr	Pro Val 1325 Ile	Ser 1310 Gln	1295 Leu) Asn	Glu Asn Thr
Glu Arg Lys Gln	Ser Gln Thr Leu 1330	Thr Asp 1315 Thr	Asp 1300 Val Val	1285 Ser) Ile Glu	Pro Lys His Thr	Asp Lys Gly Thr 1335	Glu Gln 1320 Asn	Glu 1305 Asp) Ile	1290 Thr b Lys Glu	Ser) Ile Ser Gly	Lys Ser Asp Thr 1340	Pro Val 1325 Ile	Ser 1310 Gln Ser	1295 Leu) Asn Tyr	Glu Asn Thr
Glu Arg Lys Gln	Ser Gln Thr Leu 1330	Thr Asp 1315 Thr	Asp 1300 Val	1285 Ser) Ile Glu	Pro Lys His Thr	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn	Glu 1305 Asp) Ile	1290 Thr b Lys Glu	Ser Ile Ser Gly	Lys Ser Asp Thr 1340 Asp	Pro Val 1325 Ile	Ser 1310 Gln Ser	1295 Leu) Asn Tyr	Glu Asn Thr Pro
Glu Arg Lys Gln Leu 1345	Ser Gln Thr Leu 1330 Glu	Thr Asp 1315 Thr Glu	Asp 1300 Val Val Thr	1285 Ser) Ile Glu Lys	Pro Lys His Thr Ile 1350	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn Arg	Glu 1305 Asp) Ile Tyr	1290 Thr Dys Glu Phe	Ser Ile Ser Gly Pro 1355	Lys Ser Asp Thr 1340 Asp	Pro Val 1325 Ile) Glu	Ser 1310 Gln Ser	1295 Leu) Asn Tyr	Glu Asn Thr Pro Asn 136
Glu Arg Lys Gln Leu 1345	Ser Gln Thr Leu 1330 Glu	Thr Asp 1315 Thr Glu	Asp 1300 Val Val Thr	1285 Ser) Ile Glu Lys	Pro Lys His Thr Ile 1350	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn Arg	Glu 1305 Asp) Ile Tyr	1290 Thr Dys Glu Phe	Ser Ile Ser Gly Pro 1355	Lys Ser Asp Thr 1340 Asp	Pro Val 1325 Ile) Glu	Ser 1310 Gln Ser	1295 Leu) Asn Tyr	Glu Asn Thr Pro Asn 136
Glu Arg Lys Gln Leu 1345 Ala	Ser Gln Thr Leu 1330 Glu Cys	Thr Asp 1315 Thr Glu Lys	Asp 1300 Val Val Thr	1285 Ser Ile Glu Lys Met 1365	Pro Lys His Thr Ile 1350 Lys	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn Arg	Glu 1305 Asp) Ile Tyr	1290 Thr Lys Glu Phe Phe	Ser Ile Ser Gly Pro 1355 Val	Lys Ser Asp Thr 1340 Asp Tyr	Pro Val 1325 Ile) Glu Ser	Ser 1310 Gln Ser Thr	1295 Leu) Asn Tyr Ile Gly 1375	Glu Asn Thr Pro Asn 136 Arg
Glu Arg Lys Gln Leu 1345 Ala	Ser Gln Thr Leu 1330 Glu Cys	Thr Asp 1315 Thr Glu Lys	Asp 1300 Val Val Thr	1285 Ser Ile Glu Lys Met 1365	Pro Lys His Thr Ile 1350 Lys	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn Arg	Glu 1305 Asp) Ile Tyr	1290 Thr Lys Glu Phe Phe	Ser Ile Ser Gly Pro 1355 Val	Lys Ser Asp Thr 1340 Asp Tyr	Pro Val 1325 Ile) Glu Ser	Ser 1310 Gln Ser Thr	1295 Leu) Asn Tyr Ile Gly 1375	Glu Asn Thr Pro Asn 136 Arg
Glu Arg Lys Gln Leu 1345 Ala	Ser Gln Thr Leu 1330 Glu Cys	Thr Asp 1315 Thr Glu Lys	Asp 1300 Val Val Thr	1285 Ser Ile Glu Lys Met 1365 Gly	Pro Lys His Thr Ile 1350 Lys	Asp Lys Gly Thr 1335 Thr	Glu Gln 1320 Asn Arg	Glu 1305 Asp) Ile Tyr Ser	1290 Thr Lys Glu Phe Phe 1370 Val	Ser Ile Ser Gly Pro 1355 Val	Lys Ser Asp Thr 1340 Asp Tyr	Pro Val 1325 Ile) Glu Ser	Ser 1310 Gln Ser Thr	1295 Leu Asn Tyr Ile Gly 1375 Ile	Glu Asn Thr Pro Asn 136 Arg
Arg Lys Gln Leu 1345 Ala	Gln Thr Leu 1330 Glu Cys Leu	Asp 1315 Thr Glu Lys Val	Asp 1300 Val Val Thr Thr Gly 1380	1285 Ser Ile Glu Lys Met 1365 Gly	Pro Lys His Thr Ile 1350 Lys Val	Asp Lys Gly Thr 1335 Thr Ser Asn	Glu Gln 1320 Asn Arg Arg	Glu 1305 Asp) Ile Tyr Ser Asp 1385	1290 Thr Lys Glu Phe Phe 1370 Val	Ser Ile Ser Gly Pro 1355 Val Glu	Lys Ser Asp Thr 1340 Asp Tyr	Val 1325 Ile Glu Ser	Ser 1310 Gln Ser Thr Arg Ser 1390	1295 Leu Asn Tyr Ile Gly 1375 Ile	Glu Asn Thr Pro Asn 136 Arg
Arg Lys Gln Leu 1345 Ala	Gln Thr Leu 1330 Glu Cys Leu	Asp 1315 Thr Glu Lys Val	Asp 1300 Val Val Thr Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly	Pro Lys His Thr Ile 1350 Lys Val	Asp Lys Gly Thr 1335 Thr Ser Asn Glu	Glu Gln 1320 Asn Arg Arg Gln Trp	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln	1290 Thr Lys Glu Phe Phe 1370 Val	Ser Ile Ser Gly Pro 1355 Val Glu	Lys Ser Asp Thr 1340 Asp Tyr	Val 1325 Ile Glu Ser Ser	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys	1295 Leu Asn Tyr Ile Gly 1375 Ile	Glu Asn Thr Pro Asn 136 Arg
Glu Arg Lys Gln Leu 1345 Ala Lys Asp	Gln Thr Leu 1330 Glu Cys Leu Gln	Asp 1315 Thr Glu Lys Val Gln 1395	Asp 1300 Val Val Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly	Pro Lys His Thr Ile 1350 Lys Val	Asp Lys Gly Thr 1335 Thr Ser Asn Glu	Glu Gln 1320 Asn Arg Arg Gln Trp 1400	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln	1290 Thr Lys Glu Phe 1370 Val	Ser Ile Ser Gly Pro 1355 Val Glu Gln	Lys Ser Asp Thr 1340 Asp Tyr Tyr	Val 1325 Ile Glu Ser Ser Gln 1405	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys	1295 Leu Asn Tyr Ile Gly 1375 Ile Ile	Asn Thr Pro Asn 136 Arg
Glu Arg Lys Gln Leu 1345 Ala Lys Asp	Ser Gln Thr Leu 1330 Glu Cys Leu Gln Ser	Thr Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly	Pro Lys His Thr Ile 1350 Lys Val	Asp Lys Gly Thr 1335 Thr Ser Asn Glu	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln	1290 Thr Lys Glu Phe 1370 Val	Ser Ile Ser Gly Pro 1355 Val Glu Gln	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser	Val 1325 Ile Glu Ser Ser Gln 1405 Glu	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys	1295 Leu Asn Tyr Ile Gly 1375 Ile Ile	Asn Thr Pro Asn 136 Arg
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410	Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly Thr	Pro Lys His Thr Ile 1350 Lys Val Thr	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln	1290 Thr Lys Glu Phe 1370 Val Cys	Ser) Ile Ser Gly Pro 1355 Val) Glu Gln Val	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420	Val 1325 Ile Glu Ser Ser Gln 1405 Glu	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys	1295 Leu Asn Tyr Ile Gly 1375 Ile Ile Ala	Asn Thr Pro Asn 136 Arg Thr Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410	Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly Thr	Pro Lys His Thr Ile 1350 Lys Val Thr	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln	1290 Thr Lys Glu Phe 1370 Val Cys	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln	Val 1325 Ile Glu Ser Ser Gln 1405 Glu	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys	1295 Leu Asn Tyr Ile Gly 1375 Ile Ile Ala	Glu Asn Thr Pro Asn 136 Arg Thr Thr Val
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala	Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Gly 1380 Leu Ser	1285 Ser Ile Glu Lys Met 1365 Gly Thr	Pro Lys His Thr Ile 1350 Lys Val Thr Asp Glu 1430	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln	Glu Gln 1320 Asn Arg Gln Trp 1400 Phe	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala	1295 Leu) Asn Tyr Ile Gly 1375 Ile) Ile Ala	Asn Thr Pro Asn 136 Arg Thr Thr Val
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala	Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Gly 1380 Leu	1285 Ser Ile Glu Lys Met 1365 Gly Thr	Pro Lys His Thr Ile 1350 Lys Val Thr Asp Glu 1430	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln	Glu Gln 1320 Asn Arg Gln Trp 1400 Phe	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala	1295 Leu) Asn Tyr Ile Gly 1375 Ile) Ile Ala	Asn Thr Pro Asn 136 Arg Thr Thr Val
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala	Asp 1315 Thr Glu Lys Val Gln 1395 His	Asp 1300 Val Val Thr Gly 1380 Leu Ser	1285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys	Pro Lys His Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln	Glu Gln 1320 Asn Arg Gln Trp 1400 Phe	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala	1295 Leu) Asn Tyr Ile Gly 1375 Ile) Ile Ala	Glu Asn Thr Pro Asn 136 Arg Thr Thr Val His 144 Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu	Asp 1300 Val Thr Thr Gly 1380 Leu Ser Gln	1285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445	Pro Lys His Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln Arg	Glu Gln 1320 Asn Arg Gln Trp 1400 Pro Phe	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp 1450	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Ser	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala	I285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys His Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln Arg	Glu Gln 1320 Asn Arg Gln Trp 1400 Pro Phe	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Ser	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val	Gln Leu 1330 Glu Cys Leu Gln Ser 1410 Ala Arg	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu Glu Asn	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala	I285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys Lys Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro Glu	Asp Lys Gly Thr 1335 Thr Ser Asn Glu . Ile 1415 Gln Arg Asn	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro Phe Ile	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu Val	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln Ser Lys	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp Leu	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Thr Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val	Gln Leu 1330 Glu Cys Leu Gln Ser 1410 Ala Arg	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu Glu Asn	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala Gly 1460 Pro	I285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys Lys Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro Glu	Asp Lys Gly Thr 1335 Thr Ser Asn Glu . Ile 1415 Gln Arg Asn	Glu Gln 1320 Asn Arg Gln Trp 1400 Pro Phe Ile Leu Ser	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro Leu 1465 Lys	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu Val	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln Ser Lys	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp Leu Pro	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr Asp 1470 His	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Thr Thr
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val Asp Leu	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala Arg Gly	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu Glu Asn Phe 1475	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala Gly 1460 Pro	1285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys Lys Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro Glu Leu	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln Arg Asn	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro Phe Ile Leu Ser 1480	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro Leu 1465 Lys	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp Arg 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu Val Leu	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln Ser Lys His	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp Leu Pro Gly 1485	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr Asp 1470 His	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln Pro	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Val His 144 Thr Thr Arg
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val Asp Leu	Gln Leu 1330 Glu Cys Leu Gln Ser 1410 Ala Arg Gly Val	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu Glu Asn Phe 1475 Lys	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala Gly 1460 Pro	1285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys Lys Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro Glu Leu	Asp Lys Gly Thr 1335 Thr Ser Asn Glu . Ile 1415 Gln Arg Arg Gly	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro Eleu Ser 1480 Lys	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro Leu 1465 Lys	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp Arg 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu Val Leu	Lys Ser Asp Thr 1340 Fr Tyr Tyr Val Ser 1420 Gln Ser Lys His Ser	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp Leu Pro Gly 1485 Gly	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr Asp 1470 His	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln Pro	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Val His 144 Thr Thr Arg
Glu Arg Lys Gln Leu 1345 Ala Lys Asp Arg Gln 1425 Val Asp Leu	Gln Thr Leu 1330 Glu Cys Leu Gln Ser 1410 Ala Arg Gly	Asp 1315 Thr Glu Lys Val Gln 1395 His Glu Glu Asn Phe 1475 Lys	Asp 1300 Val Val Thr Thr Gly 1380 Leu Ser Gln Ala Gly 1460 Pro	1285 Ser Ile Glu Lys Met 1365 Gly Thr Thr Lys Ile 1445 Met	Pro Lys Lys Thr Ile 1350 Lys Val Thr Asp Glu 1430 Pro Glu Leu	Asp Lys Gly Thr 1335 Thr Ser Asn Glu Ile 1415 Gln Arg Asn	Glu Gln 1320 Asn Arg Arg Gln Trp 1400 Pro Eleu Ser 1480 Lys	Glu 1305 Asp Ile Tyr Ser Asp 1385 Gln Tyr Ala Pro Leu 1465 Lys	1290 Thr Lys Glu Phe 1370 Val Cys Ile Asp Arg 1450 Ser	Ser Ile Ser Gly Pro 1355 Val Glu Gln Val Met 1435 Leu Val Leu	Lys Ser Asp Thr 1340 Asp Tyr Tyr Val Ser 1420 Gln Ser Lys His	Val 1325 Ile Glu Ser Ser Gln 1405 Glu Asp Leu Pro Gly 1485 Gly	Ser 1310 Gln Ser Thr Arg Ser 1390 Lys Ala Glu Thr Asp 1470 His	1295 Leu Asn Tyr Ile Gly 1375 Ile Ala His Ile 1455 Gln Pro	Glu Asn Thr Pro Asn 136 Arg Thr Thr Thr Val His 144 Thr Thr Arg

-13-

 Ser Val
 Ser Ser Leu
 Val
 Ile Val
 Ser Gly
 Met Thr
 Ala Glu
 Glu
 Lys

 1505
 1510
 1515
 152

 Lys Val
 Lys
 Glu
 Lys
 Ala
 Ser
 Thr
 Glu
 Thr
 Glu
 Cys

 1525
 1530

<210> 11 <211> 6220 <212> DNA <213> Homo Sapiens

<400> 11

tgtgcagaat tgtacagttg cgaaaccatg tcgctggcag ctggtgctgg cggtggagac 60 120 ttccctqtqc qqtqctcaqt gcatctgcac ccgtggggga gggagctctt tctctggccc tgcagtcacc tgaggttgtt accattatga acggccgctg ggacccccgc atgtgcatgt 180 240 actococcag agtgtocggg ggcoccagoo aagggacaca totcacgcag otgggaacat gtgcaggctg atgaagagaa ccggatgagg gcttcacatg aggaagcatg tggccaggtc 300 360 ctctcagaac atcagcctca tcttcctgtc tctgatctat ttcaccaacc accccatgtg tctctagaac cccagtgtag cgagctggag agaggactgt cctgagggca gcaggcctgg 420 ttgcagctgg cgtgggggtc tcagaatgga gccctcagcc ctgaggaaag ctggctcgga 480 gcaggaggag ggctttgagg ggctgcccag aagggtcact gacctgggga tggtctccaa 540 600 tctccggcgc agcaacagca gcctcttcaa gagctggagg ctacagtgcc ccttcggcaa caatgacaag caagaaagcc tcagttcgtg gattcctgaa aacatcaaga agaaagaatg 660 720 cgtgtatttt gtggaaagtt ccaaactgtc tgatgctggg aaggtggtgt gtcagtgtgg 780 ctacacgcat gagcagcact tggaggaggc taccaagccc cacaccttcc agggcacaca gtgggaccca aagaaacatg tccaggagat gccaaccgat gcctttggcg acatcgtctt 840 900 cacgggcctg agccagaagg tgaaaaagta cgtccgagtc tcccaggaca cgccctccag 960 cgtgatctac cacctcatga cccagcactg ggggctggac gtccccaatc tcttgatctc ggtgaccggg ggggccaaga acttcaacat gaagccgcgg ctgaagagca ttttccgcag 1020 aggcctggtc aaggtggctc agaccacagg ggcctggatc atcacagggg ggtcccacac 1080 1140 cggcgtcatg aagcaggtag gcgaggcggt gcgggacttc agcctgagca gcagctacaa ggaaggegag ctcatcacca teggagtege caeetgggge aetgtecaee geegegaggg 1200 1260 cctgatccat cccacgggca gcttccccgc cgagtacata ctggatgagg atggccaagg gaacctgacc tgcctagaca gcaaccactc tcacttcatc ctcgtggacg acgggaccca 1320 1380 cggccagtac ggggtggaga ttcctctgag gaccaggctg gagaagttca tatcggagca gaccaaggaa agaggaggtg tggccatcaa gatccccatc gtgtgcgtgg tgctggaggg 1440 cggcccgggc acgttgcaca ccatcgacaa cgccaccacc aacggcaccc cctgtgtggt 1500 tgtggagggc tcgggccgcg tggccgacgt cattgcccag gtggccaacc tgcctgtctc 1560 1620 ggacatcact atctccctga tccagcagaa actgagcgtg ttcttccagg agatgtttga gacetteacg gaaagcagga ttgtcgagtg gaccaaaaag atccaagata ttgtccggag 1680 1740 gcggcagctg ctgactgtct tccgggaagg caaggatggt cagcaggacg tggatgtggc catcttgcag gccttgctga aagcctcacg gagccaagac cactttggcc acgagaactg 1800 1860 ggaccaccag ctgaaactgg cagtggcatg gaatcgcgtg gacattgccc gcagtgagat cttcatggat gagtggcagt ggaagccttc agatctgcac cccacgatga cagctgcact 1920 catctccaac aagcctgagt ttgtgaagct cttcctggaa aacggggtgc agctgaagga 1980 gtttgtcacc tgggacacct tgctctacct gtacgagaac ctggacccct cctgcctgtt 2040 ccacagcaag ctgcaaaagg tgctggtgga ggatcccgag cgcccggctt gcgcgcccgc 2100 2160 ggcgccccgc ctgcagatgc accacgtggc ccaggtgctg cgggagctgc tgggggactt cacgcagccg ctttatcccc ggccccggca caacgaccgg ctgcggctcc tgctgcccgt 2220 tecceaegte aageteaaeg tgeagggagt gageeteegg teeetetaea agegtteete 2280 aggecatgtg accttcacca tggaccccat ccgtgacctt ctcatttggg ccattgtcca 2340 2400 gaaccgtcgg gagctggcag gaatcatctg ggctcagagc caggactgca tcgcagcggc cttggcctgc agcaagatcc tgaaggaact gtccaaggag gaggaggaca cggacagctc 2460 2520 ggaggagatg ctggcgctgg cggaggagta tgagcacaga gccatcgggg tcttcaccga gtgctaccgg aaggacgaag agagagccca gaaactgctc acccgcgtgt ccgaggcctg 2580 ggggaagacc acctgcctgc agctcgccct ggaggccaag gacatgaagt ttgtgtctca 2640 2700 cgggggcatc caggccttcc tgaccaaggt gtggtggggc cagctctccg tggacaatgg 2760 gctgtggcgt gtgaccctgt gcatgctggc cttcccgctg ctcctcaccg gcctcatctc 2820 cttcagggag aagaggctgc aggatgtggg cacccccgcg gcccgcgccc gtgccttctt caccgcaccc gtggtggtct tccacctgaa catcctctcc tacttcgcct tcctctgcct 2880 2940 gttcgcctac gtgctcatgg tggacttcca gcctgtgccc tcctggtgcg agtgtgccat

			-1-1-			
ctacctctgg	ctcttctcct	tggtgtgcga	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gracttcagt	gacttctgga	ataagctgga	3060
cgtcggcgca	atcttgctct	tcgtggcagg	gctgacctgc	aggctcatcc	cggcgacgct	3120
gtaccccggg	cgcgtcatcc	tctctctgga	cttcatcctg	ttctgcctcc	ggctcatgca	3180
	atcagtaaga					3240
ggacgtcttc	ttcttcctct	tcctgctggc	tgtgtgggtg	gtgtccttcg	gggtggccaa	3300
gcaggccatc	ctcatccaca	acgagcgccg	ggtggactgg	ctgttccgag	gggccgtcta	3360
ccactcctac	ctcaccatct	tcgggcagat	cccgggctac	atcgacggtg	tgaacttcaa	3420
cccggagcac	tgcagcccca	atggcaccga	cccctacaag	cctaagtgcc	ccgagagcga	3480
cgcgacgcag	cagaggccgg	ccttccctga	gtggctgacg	gtcctcctac	tctgcctcta	3540
cctgctcttc	accaacatcc	tgctgctcaa	cctcctcatc	gccatgttca	actacacctt	3600
	caggagcaca					3660
ggagtaccac	ggccgccccg	ccgcgccgcc	ccccttcatc	ctcctcagcc	acctgcagct	3720
cttcatcaag	agggtggtcc	tgaagactcc	ggccaagagg	cacaagcagc	tcaagaacaa	3780
gctggagaag	aacgaggagg	cggccctgct	atcctgggag	atctacctga	aggagaacta	3840
cctccagaac	cgacagttcc	agcaaaagca	gcggcccgag	cagaagatcg	aggacatcag	3900
caataaggtt	gacgccatgg	tggacctgct	ggacctggac	ccactgaaga	ggtcgggctc	3960
catggagcag	aggttggcct	ccctggagga	gcaggtggcc	cagacagccc	gagccctgca	4020
ctggatcgtg	aggacgctgc	gggccagcgg	cttcagctcg	gaggcggacg	tccccactct	4080
ggcctcccag	aaggccgcgg	aggagccgga	tgctgagccg	ggaggcagga	agaagacgga	4140
ggagccgggc	gacagctacc	acgtgaatgc	ccggcacctc	ctctacccca	actgccctgt	4200
cacgcgcttc	cccgtgccca	acgagaaggt	gccctgggag	acggagttcc	tgatctatga	4260
cccacccttt	tacacggcag	agaggaagga	cgcggccgcc	atggacccca	tgggagacac	4320
cctggagcca	ctgtccacga	tccagtacaa	cgtggtggat	ggcctgaggg	accgccggag	4380
cttccacggg	ccgtacacag	tgcaggccgg	gttgcccctg	aaccccatgg	gccgcacagg	4440
actgcgtggg	cgcgggagcc	tcagctgctt	cggacccaac	cacacgctgt	accccatggt	4500
cacgcggtgg	aggcggaacg	aggatggagc	catctgcagg	aagagcataa	agaagatgct	4560
	gtggtgaagc					4620
	gagatgctac					4680
gtcttttgaa	aacttgctga	agtgcggcat	ggaggtgtac	aaaggctaca	tggatgaccc	4740
gaggaacacg	gacaatgcct	ggatcgagac	ggtggccgtc	agcgtccact	tccaggacca	4800
	gagctgaaca					4860
	caggtggtgg					4920
ccagaaggca	gccgctgagt	tcggggctca	ctactgactg	tgccctcagg	ctgggcggct	4980
ccagtccata	gacgttcccc	ccagaaacca	gggcttctct	ctcctgagcc	tggccaggac	5040
tcaggctgtt	cctgggccct	gcacatgatg	gggtttggtg	gacccagtgc	ccctcacggc	5100
tgccgcaagt	ctgctgcaga	tgacctcatg	aactggaagg	ggtcaaggtg	acccgggagg	5160
	acagggcaca					5220
	gctgggcctg					5280
tgtgtcaccc	cgagcagctc	atccaccatg	gaggtcattg	gcctgaggca	agttccccgg	5340 5400
agagtcggga	tcccctgtgg	cccctcagg	cctatgtctg	tgaggaaggg	gccctgccac	
tctccccaag	agggcctcca	tgtttcgagg	tgcctcaaca	tggagccttg	cetggeetgg	5460 5520
gctaggggca	ctgtctgaac	tcctgactgt	caggataaac	tccgtggggg	tacaggagee	5520 5580
cagacaaagc	ccaggcctgt	caagagacgc	agagggcccc	tgccagggtt	ggccccaggg	
accctgggac	gaggctgcag	aagctctccc	tccctactcc	ctgggagcca	cgtgctggcc	5640
atgtggccag	ggacggcatg	agcaggaggc	ggggacgtgg	gggccttctg	gtttggtgtc	5700 5760
aacagctcac	aggagcgtga	accatgaggg	ccctcaggag	gggaacgrgg	taaaacccaa	
gacattaaat	ctgccatctc	aggcctggct	ggetettetg	igetttecae	tagatagagut	5820 5880
cctgacacgt	ccagggccag	gggctgtgtg	acggctgcct	gaagttctcc	cogarcecee	5940
ggtgagcttc	ctgcagcctg	rggargtcct	gcagcccctc	agecetacee	ccaagttttt	6000
cctctgaccc	atcagctccc	tgtcttcatt	ttcctaaacc	tgggctccag	categreece	6060
aagcccacca	ggccaggatg	caggcatcca	catgccctcc	teettggett	accetgegtg	6120
gcggtgccaa	tgtgccctgg	cacccctgca	gaggeteegg	acggageetg	tatattana	6180
	actggccgag			gacaggcctc	Lgracecac	6220
ccgacccaaa	gctctctagc	cacccccttg	coccagiat			0220

<212> PRT <213> Homo Sapiens

<400> 12 Met Glu Pro Ser Ala Leu Arg Lys Ala Gly Ser Glu Gln Glu Gly Phe Glu Gly Leu Pro Arg Arg Val Thr Asp Leu Gly Met Val Ser Asn Leu Arg Arg Ser Asn Ser Ser Leu Phe Lys Ser Trp Arg Leu Gln Cys Pro Phe Gly Asn Asn Asp Lys Gln Glu Ser Leu Ser Ser Trp Ile Pro Glu Asn Ile Lys Lys Lys Glu Cys Val Tyr Phe Val Glu Ser Ser Lys Leu Ser Asp Ala Gly Lys Val Val Cys Gln Cys Gly Tyr Thr His Glu Gln His Leu Glu Glu Ala Thr Lys Pro His Thr Phe Gln Gly Thr Gln Trp Asp Pro Lys Lys His Val Gln Glu Met Pro Thr Asp Ala Phe Gly Asp Ile Val Phe Thr Gly Leu Ser Gln Lys Val Lys Lys Tyr Val Arg Val Ser Gln Asp Thr Pro Ser Ser Val Ile Tyr His Leu Met Thr Gln His Trp Gly Leu Asp Val Pro Asn Leu Leu Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Asn Met Lys Pro Arg Leu Lys Ser Ile Phe Arg Arg Gly Leu Val Lys Val Ala Gln Thr Thr Gly Ala Trp Ile Ile Thr Gly Gly Ser His Thr Gly Val Met Lys Gln Val Gly Glu Ala Val Arg Asp Phe Ser Leu Ser Ser Ser Tyr Lys Glu Gly Glu Leu Ile Thr Ile Gly Val Ala Thr Trp Gly Thr Val His Arg Arg Glu Gly Leu Ile His Pro Thr Gly Ser Phe Pro Ala Glu Tyr Ile Leu Asp Glu Asp Gly Gln Gly Asn Leu Thr Cys Leu Asp Ser Asn His Ser His Phe Ile Leu Val Asp Asp Gly Thr His Gly Gln Tyr Gly Val Glu Ile Pro Leu Arg Thr Arg Leu Glu Lys Phe Ile Ser Glu Gln Thr Lys Glu Arg Gly Gly Val Ala Ile Lys Ile Pro Ile Val Cys Val Val Leu Glu Gly Gly Pro Gly Thr Leu His Thr Ile Asp Asn Ala Thr Thr Asn Gly Thr Pro Cys Val Val Val Glu Gly Ser Gly Arg Val Ala Asp Val Ile Ala Gln Val Ala Asn Leu Pro Val Ser Asp Ile Thr Ile Ser Leu Ile Ġln Gln Lys Leu Ser Val Phe Phe Gln Glu Met Phe Glu Thr Phe Thr Glu Ser Arg Ile Val Glu Trp Thr Lys Lys Ile Gln Asp Ile Val Arg Arg Arg Gln Leu Leu Thr Val Phe Arg Glu Gly Lys Asp Gly Gln Gln Asp Val Asp Val Ala Ile Leu Gln Ala Leu Leu Lys Ala Ser Arg Ser Gln Asp His Phe Gly

His	Glu 450	Asn	Trp	Asp	His	Gln 455	Leu	Lys	Leu	Ala	Val 460	Ala	Trp	Asn	Arg
465	-		Ala	•	470					475					480
			Leu	485					490					495	
			Val 500					505					510		
		515	Trp				520					525			
	530		Phe			535					540				
545	_		Ala		550					555					560
			Val	565					570					575	
_		_	Pro 580					585					590		
		595	Lys				600	_				605			
	610		Ser			615					620				
625			Trp		630					635					640
			Gln	645					650					655	
			Lys 660					665					670		
		675	Leu				680					685			
17 ~ T	Phe	ጥከ⊷	Glu	~	Π~					(2 I 11	Λνα				
	690			-	_	695	_	_			700			Lys	
Leu 705	690 Thr	Arg	Val	Ser	Glu 710	695 Ala	Trp	Gly	Lys	Thr 715	700 Thr	Cys	Leu	Gln	Leu 720
Leu 705 Ala	690 Thr Leu	Arg Glu	Val Ala	Ser Lys 725	Glu 710 Asp	695 Ala Met	Trp Lys	Gly Phe	Lys Val 730	Thr 715 Ser	700 Thr His	Cys Gly	Leu Gly	Gln Ile 735	Leu 720 Gln
Leu 705 Ala Ala	690 Thr Leu Phe	Arg Glu Leu	Val Ala Thr 740	Ser Lys 725 Lys	Glu 710 Asp	695 Ala Met Trp	Trp Lys Trp	Gly Phe Gly 745	Lys Val 730 Gln	Thr 715 Ser Leu	700 Thr His Ser	Cys Gly Val	Leu Gly Asp 750	Gln Ile 735 Asn	Leu 720 Gln Gly
Leu 705 Ala Ala Leu	690 Thr Leu Phe Trp	Arg Glu Leu Arg 755	Val Ala Thr 740 Val	Ser Lys 725 Lys Thr	Glu 710 Asp Val Leu	695 Ala Met Trp Cys	Trp Lys Trp Met 760	Gly Phe Gly 745 Leu	Lys Val 730 Gln Ala	Thr 715 Ser Leu Phe	700 Thr His Ser Pro	Cys Gly Val Leu 765	Leu Gly Asp 750 Leu	Gln Ile 735 Asn Leu	Leu 720 Gln Gly Thr
Leu 705 Ala Ala Leu Gly	690 Thr Leu Phe Trp Leu 770	Arg Glu Leu Arg 755 Ile	Val Ala Thr 740 Val Ser	Ser Lys 725 Lys Thr	Glu 710 Asp Val Leu Arg	695 Ala Met Trp Cys Glu 775	Trp Lys Trp Met 760 Lys	Gly Phe Gly 745 Leu Arg	Lys Val 730 Gln Ala Leu	Thr 715 Ser Leu Phe Gln	700 Thr His Ser Pro Asp 780	Cys Gly Val Leu 765 Val	Leu Gly Asp 750 Leu Gly	Gln Ile 735 Asn Leu Thr	Leu 720 Gln Gly Thr
Leu 705 Ala Ala Leu Gly Ala 785	690 Thr Leu Phe Trp Leu 770 Ala	Arg Glu Leu Arg 755 Ile Arg	Val Ala Thr 740 Val Ser Ala	Ser Lys 725 Lys Thr Phe Arg	Glu 710 Asp Val Leu Arg Ala 790	695 Ala Met Trp Cys Glu 775 Phe	Trp Lys Trp Met 760 Lys Phe	Gly Phe Gly 745 Leu Arg	Lys Val 730 Gln Ala Leu Ala	Thr 715 Ser Leu Phe Gln Pro 795	700 Thr His Ser Pro Asp 780 Val	Cys Gly Val Leu 765 Val	Leu Gly Asp 750 Leu Gly Val	Gln Ile 735 Asn Leu Thr	Leu 720 Gln Gly Thr Pro His 800
Leu 705 Ala Ala Leu Gly Ala 785 Leu	690 Thr Leu Phe Trp Leu 770 Ala	Arg Glu Leu Arg 755 Ile Arg .	Val Ala Thr 740 Val Ser Ala Leu	Ser Lys 725 Lys Thr Phe Arg Ser 805	Glu 710 Asp Val Leu Arg Ala 790 Tyr	695 Ala Met Trp Cys Glu 775 Phe	Trp Lys Trp Met 760 Lys Phe	Gly Phe Gly 745 Leu Arg Thr	Lys Val 730 Gln Ala Leu Ala Leu 810	Thr 715 Ser Leu Phe Gln Pro 795 Cys	700 Thr His Ser Pro Asp 780 Val Leu	Cys Gly Val Leu 765 Val Val	Leu Gly Asp 750 Leu Gly Val	Gln Ile 735 Asn Leu Thr Phe Tyr 815	Leu 720 Gln Gly Thr Pro His 800 Val
Leu 705 Ala Ala Leu Gly Ala 785 Leu	690 Thr Leu Phe Trp Leu 770 Ala Asn	Arg Glu Leu Arg 755 Ile Arg . Ile Val	Val Ala Thr 740 Val Ser Ala Leu Asp 820	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe	Glu 710 Asp Val Leu Arg Ala 790 Tyr	695 Ala Met Trp Cys Glu 775 Phe Phe	Trp Lys Trp Met 760 Lys Phe Ala Val	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825	Lys Val 730 Gln Ala Leu Ala Leu 810 Ser	Thr 715 Ser Leu Phe Gln Pro 795 Cys	700 Thr His Ser Pro Asp 780 Val Leu Cys	Cys Gly Val Leu 765 Val Val Phe	Leu Gly Asp 750 Leu Gly Val Ala Cys 830	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala	Leu 720 Gln Gly Thr Pro His 800 Val
Leu 705 Ala Ala Leu Gly Ala 785 Leu Leu	690 Thr Leu Phe Trp Leu 770 Ala Asn Met Leu	Arg Glu Leu Arg 755 Ile Arg . Ile Val Trp 835	Val Ala Thr 740 Val Ser Ala Leu Asp 820 Leu	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe	Glu 710 Asp Val Leu Arg Ala 790 Tyr Gln Ser	695 Ala Met Trp Cys Glu 775 Phe Phe	Trp Lys Trp Met 760 Lys Phe Ala Val Val 840	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825 Cys	Lys Val 730 Gln Ala Leu Ala Leu Ser Glu	Thr 715 Ser Leu Phe Gln Pro 795 Cys Trp	700 Thr His Ser Pro Asp 780 Val Leu Cys Met	Cys Gly Val Leu 765 Val Val Phe Glu Arg 845	Leu Gly Asp 750 Leu Gly Val Ala Cys 830 Gln	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala Leu	Leu 720 Gln Gly Thr Pro His 800 Val Ile
Leu 705 Ala Ala Leu Gly Ala 785 Leu Leu Tyr	690 Thr Leu Phe Trp Leu 770 Ala Asn Met Leu Asp 850	Arg Glu Leu Arg 755 Ile Arg . Ile Val Trp 835 Pro	Val Ala Thr 740 Val Ser Ala Leu Asp 820 Leu Asp	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe Phe Glu	Glu 710 Asp Val Leu Arg Ala 790 Tyr Gln Ser Cys	695 Ala Met Trp Cys Glu 775 Phe Pro Leu Gly 855	Trp Lys Trp Met 760 Lys Phe Ala Val Val 840 Leu	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825 Cys	Lys Val 730 Gln Ala Leu Ala Leu 810 Ser Glu Lys	Thr 715 Ser Leu Phe Gln Pro 795 Cys Trp Glu Lys	700 Thr His Ser Pro Asp 780 Val Leu Cys Met Ala 860	Cys Gly Val Leu 765 Val Val Phe Glu Arg 845 Ala	Leu Gly Asp 750 Leu Gly Val Ala Cys 830 Gln Leu	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala Leu Tyr	Leu 720 Gln Gly Thr Pro His 800 Val Ile Phe
Leu 705 Ala Ala Leu Gly Ala 785 Leu Leu Tyr Tyr	690 Thr Leu Phe Trp Leu 770 Ala Asn Met Leu Asp 850 Asp	Arg Glu Leu Arg 755 Ile Arg . Ile Val Trp 835 Pro	Val Ala Thr 740 Val Ser Ala Leu Asp 820 Leu Asp	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe Phe Glu Asn	Glu 710 Asp Val Leu Arg Ala 790 Tyr Gln Ser Cys Lys 870	695 Ala Met Trp Cys Glu 775 Phe Pro Leu Gly 855 Leu	Trp Lys Trp Met 760 Lys Phe Ala Val 840 Leu Asp	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825 Cys Met Val	Lys Val 730 Gln Ala Leu 810 Ser Glu Lys Gly	Thr 715 Ser Leu Phe Gln Pro 795 Cys Trp Glu Lys Ala 875	700 Thr His Ser Pro Asp 780 Val Leu Cys Met Ala 860 Ile	Cys Gly Val Leu 765 Val Val Phe Glu Arg 845 Ala	Leu Gly Asp 750 Leu Gly Val Ala Cys 830 Gln Leu Leu	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala Leu Tyr	Leu 720 Gln Gly Thr Pro His 800 Val Ile Phe Phe Val 880
Leu 705 Ala Ala Leu Gly Ala 785 Leu Leu Tyr Tyr	690 Thr Leu Phe Trp Leu 770 Ala Asn Met Leu Asp 850 Asp	Arg Glu Leu Arg 755 Ile Arg . Ile Val Trp 835 Pro	Val Ala Thr 740 Val Ser Ala Leu Asp 820 Leu Asp	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe Phe Glu Asn	Glu 710 Asp Val Leu Arg Ala 790 Tyr Gln Ser Cys Lys 870	695 Ala Met Trp Cys Glu 775 Phe Pro Leu Gly 855 Leu	Trp Lys Trp Met 760 Lys Phe Ala Val 840 Leu Asp	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825 Cys Met Val	Lys Val 730 Gln Ala Leu 810 Ser Glu Lys Gly	Thr 715 Ser Leu Phe Gln Pro 795 Cys Trp Glu Lys Ala 875	700 Thr His Ser Pro Asp 780 Val Leu Cys Met Ala 860 Ile	Cys Gly Val Leu 765 Val Val Phe Glu Arg 845 Ala	Leu Gly Asp 750 Leu Gly Val Ala Cys 830 Gln Leu Leu	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala Leu Tyr	Leu 720 Gln Gly Thr Pro His 800 Val Ile Phe Phe Val 880
Leu 705 Ala Ala Leu Gly Ala 785 Leu Tyr Tyr Ser 865 Ala Val	690 Thr Leu Phe Trp Leu 770 Ala Asn Met Leu Asp 850 Asp Gly Ile	Arg Glu Leu Arg 755 Ile Arg . Ile Val Trp 835 Pro Phe Leu Leu	Val Ala Thr 740 Val Ser Ala Leu Asp 820 Leu Asp	Ser Lys 725 Lys Thr Phe Arg Ser 805 Phe Glu Asn Cys 885 Leu	Glu 710 Asp Val Leu Arg Ala 790 Tyr Gln Ser Cys Lys 870 Arg Asp	695 Ala Met Trp Cys Glu 775 Phe Pro Leu Gly 855 Leu Leu	Trp Lys Trp Met 760 Lys Phe Ala Val 840 Leu Asp Ile	Gly Phe Gly 745 Leu Arg Thr Phe Pro 825 Cys Met Val Pro Leu 905	Lys Val 730 Gln Ala Leu Ala Leu Sl0 Ser Glu Lys Gly Ala 890 Phe	Thr 715 Ser Leu Phe Gln Pro 795 Cys Trp Glu Lys Ala 875 Thr	700 Thr His Ser Pro Asp 780 Val Leu Cys Met Ala 860 Ile Leu Leu	Cys Gly Val Leu 765 Val Val Phe Glu Arg 845 Ala Leu Tyr	Leu Gly Asp 750 Leu Gly Val Ala Cys 830 Gln Leu Leu Pro Leu 910	Gln Ile 735 Asn Leu Thr Phe Tyr 815 Ala Leu Tyr Phe Gly 895 Met	Leu 720 Gln Gly Thr Pro His 800 Val Ile Phe Val 880 Arg

-17-

Arg Met Met Lys Asp Val Phe Phe Phe Leu Phe Leu Ala Val Trp 940 935 930 Val Val Ser Phe Gly Val Ala Lys Gln Ala Ile Leu Ile His Asn Glu 950 955 Arg Arg Val Asp Trp Leu Phe Arg Gly Ala Val Tyr His Ser Tyr Leu 975 965 970 Thr Ile Phe Gly Gln Ile Pro Gly Tyr Ile Asp Gly Val Asn Phe Asn 980 985 990 Pro Glu His Cys Ser Pro Asn Gly Thr Asp Pro Tyr Lys Pro Lys Cys 1000 1005 Pro Glu Ser Asp Ala Thr Gln Gln Arg Pro Ala Phe Pro Glu Trp Leu 1020 1015 Thr Val Leu Leu Cys Leu Tyr Leu Leu Phe Thr Asn Ile Leu Leu 1035 1030 Leu Asn Leu Leu Ile Ala Met Phe Asn Tyr Thr Phe Gln Gln Val Gln 1045 1050 Glu His Thr Asp Gln Ile Trp Lys Phe Gln Arg His Asp Leu Ile Glu 1060 1065 1070 Glu Tyr His Gly Arg Pro Ala Ala Pro Pro Pro Phe Ile Leu Leu Ser 1075 1080 1085 His Leu Gln Leu Phe Ile Lys Arg Val Val Leu Lys Thr Pro Ala Lys 1090 1095 1100 Arg His Lys Gln Leu Lys Asn Lys Leu Glu Lys Asn Glu Glu Ala Ala 1110 1115 Leu Leu Ser Trp Glu Ile Tyr Leu Lys Glu Asn Tyr Leu Gln Asn Arg 1125 1130 1135 Gln Phe Gln Gln Lys Gln Arg Pro Glu Gln Lys Ile Glu Asp Ile Ser 1140 1145 1150 Asn Lys Val Asp Ala Met Val Asp Leu Asp Leu Asp Pro Leu Lys 1155 1160 1165 Arg Ser Gly Ser Met Glu Gln Arg Leu Ala Ser Leu Glu Glu Gln Val 1175 1180 1170 Ala Gln Thr Ala Arg Ala Leu His Trp Ile Val Arg Thr Leu Arg Ala 1190 . 1195 Ser Gly Phe Ser Ser Glu Ala Asp Val Pro Thr Leu Ala Ser Gln Lys 1205 1210 1215 Ala Ala Glu Glu Pro Asp Ala Glu Pro Gly Gly Arg Lys Lys Thr Glu 1225 1230 1220 Glu Pro Gly Asp Ser Tyr His Val Asn Ala Arg His Leu Leu Tyr Pro 1235 1240 1245 Asn Cys Pro Val Thr Arg Phe Pro Val Pro Asn Glu Lys Val Pro Trp 1250 1260 Glu Thr Glu Phe Leu Ile Tyr Asp Pro Pro Phe Tyr Thr Ala Glu Arg 1275 1270 Lys Asp Ala Ala Ala Met Asp Pro Met Gly Asp Thr Leu Glu Pro Leu 1290 1295 1285 Ser Thr Ile Gln Tyr Asn Val Val Asp Gly Leu Arg Asp Arg Arg Ser 1300 1305 1310 Phe His Gly Pro Tyr Thr Val Gln Ala Gly Leu Pro Leu Asn Pro Met 1325 1320 1315 Gly Arg Thr Gly Leu Arg Gly Arg Gly Ser Leu Ser Cys Phe Gly Pro 1330 1340 1335 1340 1330 Asn His Thr Leu Tyr Pro Met Val Thr Arg Trp Arg Arg Asn Glu Asp 1355 1350 Gly Ala Ile Cys Arg Lys Ser Ile Lys Lys Met Leu Glu Val Leu Val 1370 1375 1365 Val Lys Leu Pro Leu Ser Glu His Trp Ala Leu Pro Gly Gly Ser Arg 1380 1385 Glu Pro Gly Glu Met Leu Pro Arg Lys Leu Lys Arg Ile Leu Arg Gln 1400 1405 1395

PCT/US99/29996 WO 00/40614

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val 1415 1420 Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile 1430 1435 Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu 1445 1450 1455 Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser 1465 1460 1470 Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His 1475 1480 1485 Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr 1495 1500

<210> 13 <211> 1816 <212> PRT

<213> C. Elegans

<400> 13

Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn 5 10 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser 20 25 30 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr 40 4.5 35 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu 55 60 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser 75 70 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu 85 90 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg 105 110 100 Glu Thr Glu Thr Gln Thr Lys Arg Arg Arg Lys Lys Gln Arg Ser Thr 125 115 120 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys 135 140 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys 150 155 Arg Gly Asn Ser Thr Thr Ser His Ser Gly His Ala Thr Arg Ala Gly 170 165 175 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln 180 185 190 Cys Ser Tyr Phe Val Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys 200 205 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln 215 220 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu

Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asn Glu Val Asn Leu Thr
245 250 255 . 250 245 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp 260 265 270 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr 280 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met 295 300 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn 305 310 315 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe

235

230

-19-

				225					330					335	
_		61	T 1 =	325	T	т1 о	ת דת	t uc		Thr) en	Δla	Trn	Ile	Tlo
-	-	_	34.0					345					350		
		355					360					365		Ala	
His	Ala 370	Leu	Glu	Phe	Trp	Ser 375	Phe	Gly	Leu	Phe	Trp 380	Val	Ile	Gln	Leu
Asp 385		Leu	Leu	Ala	His 390	Ser	Met	Phe	Ile	Pro 395	Arg	Gly	Ser	Leu	Phe 400
Asp	His	Gly	Asn	His 405		Ser	Lys	Asn	His 410	Vaľ	Val	Ala	Ile	Gly 415	Ile
Ala	Ser	Trp	Gly 420		Leu	Lys	Gln	Arg 425		Arg	Phe	Val	Gly 430	Lys	Asp
Ser	Thr	Val 435		Tyr	Ala	Thr	Asn 440		Phe	Asn	Asn	Thr 445	Arg	Leu	Lys
Glu	Leu 450		Asp	Asn	His	Ser 455		Phe	Leu	Phe	Ser 460	Asp	Asn	Gly	Thr
Val 465	Asn	Arg	Tyr	Gly	Ala 470		Ile	Ile	Met	Arg 475	Lys	Arg	Leu	Glu	Ala 480
Tyr	Leu	Ala	Gln	Gly 485		Lys	Lys	Arg	Ser 490	Ala	Ile	Pro	Leu	Val 495	Cys
Val	Val	Leu	Glu 500	Gly	Gly	Ala	Phe	Thr 505	Ile	Lys	Met	Val	His 510	Asp	Tyr
Val	Thr	Thr 515	Ile	Pro	Arg	Ile	Pro 520	Val	Ile	Val	Cys	Asp 525	Gly	Ser	Gly
Arg	Ala 530	Ala	Asp	Ile	Leu	Ala 535	Phe	Ala	His	Gln	Ala 540	Val	Ser	Gln	Asn
545	Phe				550					555				Val	560
				565					570					Lys 575	
			580					585					590	Arg	
		595					600					605		Cys	
	610	_				615					620			Ala	
625					630					635				Asn	640
				645					650					Ala 655	
			660					665					670	Gly	
		675					680					685		Tyr	
	690					695					700			Val	
705					710					/15				Glu	120
				725					730					Phe 735	
			740					745					750	Arg	
		755					760					765		Ser	
-	770					775					780			Ile	
785					790					795				Arg	800
Arg	Glu	Ser	Asp	Asp	Glu	Asp	Asp	Phe	Ser	Asn	Leu	Glu	Glu	Glu	Ala

-20-805 810 Asn Met Asp Phe Thr Phe Arg Tyr Pro Tyr Ser Asp Leu Met Ile Trp 820 825 Ala Val Leu Thr Lys Arg Gln Lys Met Ala Lys Leu Met Trp Thr His 840 845 Gly Glu Glu Gly Met Ala Lys Ala Leu Val Ala Ser Arg Leu Tyr Val 855 860 Ser Leu Ala Lys Thr Ala Ser Leu Ala Thr Gly Glu Ile Gly Met Ser 870 875 Gln Asp Phe Thr Glu Phe Ser Asp Glu Phe Ser Glu Leu Ala Val Glu 885 890 895 Val Leu Glu Tyr Cys Thr Lys His Gly Arg Asp Gln Thr Leu Arg Leu 900 905 910 Leu Thr Cys Glu Leu Ala Asn Trp Gly Asp Glu Thr Cys Leu Ser Leu 920 925 915 Ala Ala Asn Asn Gly His Arg Lys Phe Leu Ala His Pro Cys Cys Gln 935 930 940 Met Leu Leu Ser Asp Leu Trp Gln Gly Gly Leu Leu Met Lys Asn Asn 950 955 Gln Asn Ser Lys Val Leu Thr Cys Leu Ala Ala Pro Pro Leu lle Phe 970 975 965 Leu Leu Gly Phe Lys Thr Lys Glu Gln Leu Met Leu Gln Pro Lys Thr 990 980 985 Ala Ala Glu His Asp Glu Glu Met Ser Asp Ser Glu Met Asn Ser Ala 1000 995 1005 Glu Asp Thr Asp Thr Ser Ser Asp Ser Ser Ser Asp Ser Asp Ser 1010 1015 1020 Asp Glu Glu Asp Ala Lys Leu Arg Ala Gln Ser Leu Ser Ala Asp Gln 1025 1030 1035 Pro Leu Ser Ile His Arg Leu Val Arg Asp Lys Leu Asn Phe Ser Glu 1045 1050 1055 Lys Lys Pro Asp Met Gly Ile Ser Arg Ile Val Val Ala Pro Pro 1060 1065 1070 Ile Val Thr Gly Arg Asn Arg Ala Arg Thr Met Ser Ile Lys Lys Ser 1075 1080 1085 Lys Lys Asn Val Ile Lys Pro Pro Ala Cys Leu Lys Ile Glu Thr Ser 1090 1095 1100 Asp Asp Asp Glu Gln Glu Gln Lys Lys Ala Thr Glu Met Cys Lys Ser 1110 1115 Thr Phe Phe Asp Phe Phe Asp Phe Pro Tyr Ile Asn Arg Thr Gly 1125 1130 1135 Lys Arg Gly Ser Val Ala Val Ala Met Asn His Asp Asp Met Tyr Ile 1145 1150 1140 Asp Pro Ser Glu Glu Leu Asp Thr Gln Thr Arg Gln Lys Ser Ser Arg 1155 1160 1165 Glu Phe Ser Ser Ser Arg Asn Val Thr Val Gln Val Tyr Thr Gln Arg 1175 1180 Pro Leu Ser Trp Lys Lys Lys Ile Met Glu Phe Tyr Lys Ala Pro Ile 1190 1195 Thr Thr Tyr Trp Leu Trp Phe Phe Ala Phe Ile Trp Phe Leu Ile Leu 1205 1210 1215 Leu Thr Tyr Asn Leu Leu Val Lys Thr Gln Arg Ile Ala Ser Trp Ser 1220 1225 1230Glu Trp Tyr Val Phe Ala Tyr Ile Phe Val Trp Thr Leu Glu Ile Gly 1240 Arg Lys Val Val Ser Thr Ile Met Met Asp Thr Ser Lys Pro Val Leu 1250 1255 1260 Lys Gln Leu Arg Val Phe Phe Phe Gln Tyr Arg Asn Gly Leu Leu Ala 1270 1275

Phe Gly Leu Leu Thr Tyr Leu Ile Ala Tyr Phe Ile Arg Leu Ser Pro

				1005	_				1290	`				1295	
		_		1285) 	_	~ 3 =	.	1230	, 	C	7.00	S		
			1300)				1305)				1310		
•		1319	5				1320)				1325	5	Leu	
Pro	Tyr 1330	Ile	Asn	Ile	Val	Ala 1335		Met	Ile	Pro	Thr 1340	Met)	Ile	Pro	Leu
	Val	Leu	Val	Phe		Thr	Leu	Tyr	Ala	Phe 1355	Gly		Leu	Arg	Gln 136
1345	71-	m 1	m	D==	1350	, -	Λες	Trn	Hic			Len	Val	Arg	
Ser	11e	inr	Tyr	1365		GIU	ASD	тър	1370)	110	пса	• • • •	1375	5
			1380	Pro	Tyr			1385	Tyr	Gly			1390		
		1399	Thr	Cys			1400)				1405	5	Asp	
Asn	Ile 1410	Pro	Ile	Ser	Met	Leu 1415	Asn	Val	Thr	His	Glu 1420	Thr	Cys	Val	Pro
Glv	Tyr	Tro	Ile	Ala	Pro	Val	Gly	Leu	Thr	Val	Phe	Met	Leu	Ala	Thr
1425	5 ⁻				1430)				1435	5				144
Asn	Val	Leu	Leu	Met	Asn	Val	Met	Val	Ala	Gly	Cys	Thr	Tyr	Ile	Phe
				1445	5				1450		_		۵,	1455	
			1460	ו				1465	5				1470		
		1475	5				1480)				1483	•	Pro	
	1490)				1499	5				1500)		Ser	
Ser	Arg	Met	Phe	Glu			Asn	Leu	Phe	Asp	Gln	Ser	Leu	Lys	Leu
150	5				1510)		_		151		D1	63	C1	152
				1525	5				1530)				Glu 1535	כ
			1540)				1545	5				1550		
		155	5				1560)				156	5	Ile	
	1570)				157	5				158	0		Glu	
158	5				1590)				159	5			His	160
				1.60	5				1610)				Lys 161	5
			1620)				162)				163		
		1631	5				1640	0				164	5	Asn	
Val	Pro	Met	Ile	Thr	Val	Asp	Gly	Pro	Ser	Pro	Ile	Gly	Ser	Arg	Arg
Thr	1650 Ser) Gly	Gln	Tyr	Leu	165 Lys		Asp	Ser	Leu	166 Gln		Lys	Lys	Lys
166	5				1670)				167	5				168
				168	5				1690	0				Pro 169	5
			1700)				170	5				171		
		171	Glu 5	Val			172	0				172	5	Arg	
	1730	Val	Thr			173.	5				174	0		Asp	
174	Thr 5	Arg			1750	C				175	5			Arg	176
Pro	Arg	Arg	His	Ala	Leu	Tyr	Ser	Thr	Ile	Ala	Asp	Ala	Ile	Glu	Thr

PCT/US99/29996

WO 00/40614 -22-1770 1765 Glu Asp Asp Phe Tyr Ala Asp Ser Pro Val Pro Met Pro Met Thr Pro 1780 1785 1790 Val Gln Pro Ala Asp Gly Ser Phe Phe Gly Glu Asn Asp Ser Arg Tyr 1795 1800 1805 Gln Arg Asp Asp Ser Asp Tyr Glu <210> 14 <211> 1387 <212> PRT <213> C. Elegans <400> 14 Met Arg Lys Ser Arg Arg Val Arg Lys Leu Val Arg His Ala Ser Leu 10 - 5 Ile Glu Asn Ile Arg His Arg Thr Ser Ser Phe Leu Arg Leu Leu Asn 25 20 Ala Pro Arg Asn Ser Met Cys Asn Ala Asn Thr Val His Ser Ile Ser 45 40 35 Ser Phe Arg Ser Asp His Leu Ser Arg Lys Ser Thr His Lys Phe Leu 55 60 Asp Asn Pro Asn Leu Phe Ala Ile Glu Leu Thr Glu Lys Leu Ser Pro 75 70 Pro Trp Ile Glu Asn Thr Phe Glu Lys Arg Glu Cys Ile Arg Phe Ala 90 85 Ala Leu Pro Lys Asp Pro Glu Arg Cys Gly Cys Gly Arg Pro Leu Ser 100 105 110 Ala His Thr Pro Ala Ser Thr Phe Phe Ser Thr Leu Pro Val His Leu 125 120 115 Leu Glu Lys Glu Gln Gln Thr Trp Thr Ile Ala Asn Asn Thr Gln Thr 135 140 Ser Thr Thr Asp Ala Phe Gly Thr Ile Val Phe Gln Gly Gly Ala His 155 150 Ala His Lys Ala Gln Tyr Val Arg Leu Ser Tyr Asp Ser Glu Pro Leu 165 170 175 165 Asp Val Met Tyr Leu Met Glu Lys Val Trp Gly Leu Glu Ala Pro Arg 180 190 185 Leu Val Ile Thr Val His Gly Gly Met Ser Asn Phe Glu Leu Glu Glu 200 205 195 Arg Leu Gly Arg Leu Phe Arg Lys Gly Met Leu Lys Ala Ala Gln Thr 215 220 Thr Gly Ala Trp Ile Ile Thr Ser Gly Leu Asp Ser Gly Val Val Arg 235 230 His Val Ala Lys Ala Leu Asp Glu Ala Gly Ile Ser Ala Arg Met Arg 250 245

Ser Gln Ile Val Thr Ile Gly Ile Ala Pro Trp Gly Val Ile Lys Arg 260 265 270 Lys Glu Arg Leu Ile Arg Gln Asn Glu His Val Tyr Tyr Asp Val His 285 275 280 Ser Leu Ser Val Asn Ala Asn Val Gly Ile Leu Asn Asp Arg His Ser 295 300 Tyr Phe Leu Leu Ala Asp Asn Gly Thr Val Gly Arg Phe Gly Ala Asp 310 315 Leu His Leu Arg Gln Asn Leu Glu Asn His Ile Ala Thr Phe Gly Cys 330 335 Asn Gly Arg Lys Val Pro Val Val Cys Thr Leu Leu Glu Gly Gly Ile 345 350

Ser Ser Ile Asn Ala Ile His Asp Tyr Val Thr Met Lys Pro Asp Ile

360

PCT/US99/29996 WO 00/40614

W	O 00/4	0614													PC17
									-23-		•				
Pro	Ala 370	Ile	Val	Cys	Asp	Gly 375	Ser	Gly	Arg	Ala	Ala 380	Asp	Ile	Ile	Ser
Phe 385	Ala	Ala	Arg	Tyr	Ile 390		Ser	Asp	Gly	Thr 395		Ala	Ala	Glu	Val 400
	Glu	Lys	Leu	Arg 405		Leu	Ile	Lys	Met 410	Val	Phe	Pro	Glu	Thr 415	Asp
	Glu		420					425					430		
Leu	Leu	Arg 435	Ile	Phe	Arg	Tyr	Gly 440	Gln	Glu	Glu	Glu	Glu 445	Asp	Val	Asp
	Val 450					455			_		460				
465	Gln				470					475					480
	Cys			485					490					495	
_	Ala		500					505					510		
_	Leu -	515					520					525			
_	Leu 530					535					540				
545	Arg Thr				550					555					560
	Phe			565					570					575	
	Tyr		580					585					590		
	Val	595					600					605			
	610 Asn		-			615	_	_			620				
625	Ile				630					635					640
	Val			645					650					655	
	Tyr		660					665					670		
	Glu	675					680					685			
	690 Ile					695					700				
705 Leu	Arg	Leu	Leu	Arg	710 Met	Glu	Leu	Pro		715 Trp	Gly	Asn	Asn		720 Cys
Leu	Ser	Leu	Ala	725 Val	Leu	Ala	Asn		730 Lys	Thr	Phe	Leu		735 His	Pro
Cys	Cys		740 Ile	Leu	Leu	Ala		745 Leu	Trp	His	Gly		750 Leu	Lys	Val
Arg	Ser	755 Gly	Ser	Asn	Val		760 Val	Leu	Thr	Ala		765 Ile	Cys	Pro	Pro
	770 Ile	Leu	Phe	Met		775 Tyr	Lys	Pro	Lys		780 Ser	Lys	Thr	Ala	
785 Leu	Leu	Ser	Glu	Glu 805	790 Thr	Pro	Glu	Gln	Leu 810	795 Pro	Tyr	Pro	Arg	Glu 815	800 Ser
Ile	Thr	Ser	Thr 820		Ser	Asn	Arg	Tyr 825		Tyr	Ser	Lys	Gly 830		Glu
Glu	Gln	Lys 835		Thr	Leu	Leu	Glu 840		Gly	Ser	Tyr	Thr 845		Lys	Val

WC	00/4	0614													rC1/
									-24-						
Thr		Ile	Ser	Ser	Arg	Lys 855	Asn	Ser	Gly	Val	Ala 860	Ser	Val	Туг	Gly
	850 Ala	Ser	Ser	Met			Lys	Arg	Glu			Leu	Asn	Lys	Phe 880
865 Glu	Arg	Phe	Arg		870 Phe	Tyr	Ser	Ser		875 Ile	Thr	Lys	Phe		
Trp	Суѕ	Ile	Ala	885 Phe	Leu	Ile	Phe		890 Thr	Thr	Gln	Thr		895 Ile	Leu
Leu	Leu		900 Thr	Ser	Leu	Lys		905 Ser	Lys	Tyr	Glu	Trp 925	910 Ile	Thr	Phe
Ile	_	915 Thr	Val	Thr	Leu	Ser 935	920 Val	Glu	His	Ile	Arg 940		Leu	Met	Thr
	930 Glu	Gly	Ser	Arg	Ile 950		Glu	Lys	Val	Lys 955		Phe	Tyr	Ala	Lys 960
945 Trp	Tyr	Asn	Ile	Trp 965		Ser	Ala	Ala	Leu 970		Phe	Phe	Leu	Val 975	
Tyr	Gly	Phe	Arg 980		Val	Pro	Met	Tyr 985		His	Ser	Trp	Gly 990		Val
Leu	Leu	Ser 995	Phe	Ser	Asn	Val	Leu 1000	Phe	Tyr	Met	Lys	Ile 1009	Phe	Glu	Tyr
Leu	Ser 1010	Val	His	Pro	Leu	Leu 1015	Gly		Tyr	Ile	Gln 1020	Met		Ala	Lys
Met			Ser	Met	Cys	Tyr	Ile	Cys	Val	Leu	Leu	Leu	Val	Pro	Leu
1025	5	_			1030)				103	5				104
			Gly	1045	5				1050)				1055	5
			Trp 1060)				1069	5				1070)	
		1075					1080)				1089	5		
	1090)	Arg			1099	5				1100)			
Val 1105		Phe	Leu	Leu	Val 1110		Esn	Ile	Leu	Leu 111!		Asn	Leu	Leu	11e 112
		Phe	Asn	Asn			Asn	Asp	Ser			Lvs	Ser	Lvs	
				1125	5				1130)				1135	5
			Phe 1140)				1145	5				1150)	
		1155					1160)				1169	5		
	1170) _	Leu			1175	5				1180)			
Ser	Glu	His	Ser	Ile			Ser	Val	Thr			Glu	Met	Lys	
1185		_		٠.	1190			- 1 -	.	119		m 1	70	T1.	120
				1205	5				1210)				1215	Arg
			1220)				1225	5				1230)	Thr
		1235	5				1240)				1245	5		Leu
	1250)	Arg			1255	5				1260)			
Asn 1265		Ser	Asp	Glu	Val 1270		Gln	Ile	Leu	Lys 127		Lys	Lys	Leu	Ser 128
		Phe	Ala	Ala 1285	Ser		Leu	Ser	Leu 1290	Pro		Thr	Ser	Ile 1299	Glu
Val	Pro	Lys	Ile 1300	Thr		Thr	Leu	Ile 1305	Asp		His	Leu	Ser 1310	Pro	
Ser	Ile	Glu 1315	Asp		Leu	Ala	Thr 1320	Arg		Pro	Leu	Leu 1325	Ala		Leu

-25-

<210> 15 <211> 1868 <212> PRT <213> C. Elegans

<400> 15 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val 1 5 10 15 Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln 25 20 Ser Lys Gly Gly Asp Gln Val Pro Pro Thr Ser Thr Thr Thr Gly Gly 4.5 40 Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala 60 55 Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp 75 70 Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val 90 95 85 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met 105 110 Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu 120 125 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys 130 135 140 Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu 155 150 Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser 170 175 Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys 185 190 180 Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu 195 200 205 Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro 215 220 Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr 230 235 Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu 250 245 Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys 265 270 260 Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr 280 275 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His 300 290 295 Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile 310 315 Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Leu Lys Lys Arg Glu Asp 330 325 Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Lys 340 345 Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

-26-Asp Asn Gly Thr Val Gly Arg Tyr Gly Ala Glu Val Ile Leu Arg Lys Arg Leu Glu Met Tyr Ile Ser Gln Lys Gln Lys Ile Phe Gly Gly Thr Arg Ser Val Pro Val Val Cys Val Val Leu Glu Gly Gly Ser Cys Thr Ile Arg Ser Val Leu Asp Tyr Val Thr Asn Val Pro Arg Val Pro Val Val Val Cys Asp Gly Ser Gly Arg Ala Ala Asp Leu Leu Ala Phe Ala 435 440 445 His Gln Asn Val Thr Glu Asp Gly Leu Leu Pro Asp Asp Ile Arg Arg Gln Val Leu Leu Val Glu Thr Thr Phe Gly Cys Ser Glu Ala Ala Ala His Arg Leu Leu His Glu Leu Thr Val Cys Ala Gln His Lys Asn Leu Leu Thr Ile Phe Arg Leu Gly Glu Gln Gly Glu His Asp Val Asp His Ala Ile Leu Thr Ala Leu Leu Lys Gly Gln Asn Leu Ser Ala Ala Asp Gln Leu Ala Leu Ala Leu Ala Trp Asn Arg Val Asp Ile Ala Arg Ser Asp Val Phe Ala Met Gly His Glu Trp Pro Gln Ala Ala Leu His Asn Ala Met Met Glu Ala Leu Ile His Asp Arg Val Asp Phe Val Arg 565 570 575 Leu Leu Glu Gln Gly Ile Asn Met Gln Lys Phe Leu Thr Ile Ser Arg Leu Asp Glu Leu Tyr Asn Thr Asp Lys Gly Pro Pro Asn Thr Leu Phe Tyr Ile Val Arg Asp Val Val Arg Val Arg Gln Gly Tyr Arg Phe Lys Leu Pro Asp Ile Gly Leu Val Ile Glu Lys Leu Met Gly Asn Ser Tyr Gln Cys Ser Tyr Thr Thr Ser Glu Phe Arg Asp Lys Tyr Lys Gln Arg Met Lys Arg Val Lys His Ala Gln Lys Lys Ala Met Gly Val Phe Ser Ser Arg Pro Ser Arg Thr Gly Ser Gly Ile Ala Ser Arg Gln Ser Thr Glu Gly Met Gly Gly Val Gly Gly Gly Ser Ser Val Ala Gly Val Phe Gly Asn Ser Phe Gly Asn Gln Asp Pro Pro Leu Asp Pro His Val Asn Arg Ser Ala Leu Ser Gly Ser Arg Ala Leu Ser Asn His Ile Leu Trp Arg Ser Ala Phe Arg Gly Asn Phe Pro Ala Asn Pro Met Arg Pro Pro Asn Leu Gly Asp Ser Arg Asp Cys Gly Ser Glu Phe Asp Glu Glu Leu Ser Leu Thr Ser Ala Ser Asp Gly Ser Gln Thr Glu Pro Asp Phe

Arg Tyr Pro Tyr Ser Glu Leu Met Ile Trp Ala Val Leu Thr Lys Arg

Gln Asp Met Ala Met Cys Met Trp Gln His Gly Glu Glu Ala Met Ala

Lys Ala Leu Val Ala Cys Arg Leu Tyr Lys Ser Leu Ala Thr Glu Ala

Ala Glu Asp Tyr Leu Glu Val Glu Ile Cys Glu Glu Leu Lys Lys Tyr

	835					840					845			
Ala Glu 850	Glu	Phe	Arg	Ile	Leu 855	Ser	Leu	Glu	Leu	Leu 860	Asp	His	Cys	Tyr
His Val 865				870					875					880
Asn Trp			885					890					895	
Lys His		900					905					910		
Trp His	915	_				920					925			
Leu Gly 930			_		935					940				
Arg Glu 945				950					955					960
Asp Met		-	965					970					975	
Ser Ser		980					985					990		
Asn Asn	995				_	1000	o -	_			100	5		
Gly Ser	0				1019	5				1020)			
Arg Lys 1025		_	_	1030)				1035	5				104
Ala Cys			1045	5				1050)				105	5
Glu Tyr		1060)				106	5				1070)	
Tyr Met	107	5			_	1080	ר בי כ	-			108	5		
Ser Lys 1090)				1095	5				1100)			
Leu Gln 1105				1110)				1115	5				112
Gln Phe		_	1125	5				1130)				1135	5
Phe Tyr		1140)				114	5				1150)	
Ile Leu	1155	5				1160)				116	5		
Pro Arg)				1175	5				1180)			
Phe Gly 1185	Leu	Glu	GIn	Vai 1190		гуѕ	ııe	11e	1195		Asp	Ala	гуѕ	120
Phe Tyr		_	1205	5				1210)				1215	5
Thr Ile		1220)				122	5				1230)	
Phe Gly	1235	5				1240)				124	5		
Leu Trp 1250)				1255	5				1260)			
Gly Pro 1265				1270)				1275	5				128
Ile Ile			1285	5				1290)				1295	5
Gln Ser	Ile	Thr 1300		Pro	Asp	Glu	Thr 130		His	Trp	Ile	Leu 1310		Arg
Asn Ile	Phe			Pro	Tyr	Phe			Tyr	Gly	Glu	Val	Tyr	Ala

				-28-			
131			1320			1325	
Asp Glu Ile 1330	Asp Thr	Cys Gly 133		lu Ala	Trp Asp		His Leu Glu
	Pro Val	Ile Leu 1350	Gly A		Thr Thr 1355	Gly	Leu Ser Cys 136
	Tyr Trp 136	Ile Pro	Pro Le		Met Thr	Phe	Phe Leu Leu 1375
Ile Ala Asn						Ile	Phe Asn His
Ile Phe Asp	Ala Thr	Asp Glu			Gln Ile	Trp	Leu Phe Gln
		Met Glu 141	Tyr G	lu Ser	Thr Pro	Phe	Leu Pro Pro
	Pro Leu			al Leu			Phe Val Arg
	Ser Cys	Ser Lys	Ser G	ln Glu 1450	Arg Asr	Pro	Ile Leu Leu 1455
Leu Lys Ile						Ile	Glu Lys Leu 1470
His Asp Phe	Glu Glu	Asp Cys			Leu Ala	Arg 1485	Gln Lys Leu
	-	Ser Asn	Glu G	ln Arg	Ile Leu	Arg	Ala Asp Ile
	Gln Ile					-	Ala Lys Glu 152
	Arg Asp	Val Ile	Asn As		Glu Ser	Arg	Leu Ala Ser 1535
Val Glu Lys						Arg	Ala Leu Leu 1550
Asn Gln Asn	Asn Ala	Pro Thr			Arg Cys	Phe 1565	Ser Pro Ser
		Glu Thr	Ala As	sn Gly	Thr Pro	Gly	Pro Leu Leu
	Pro Gly						Asp His Asp
	Asn Ser 160	Asn Ser	Leu Pi		Gly Arg	Ile	Arg Arg Asn 1615
Arg Thr Ala						Glu	
Met Leu Leu 163	Ser Pro	Lys Pro			Ser Gly	Ile 1645	Pro Gln Gln
		Thr Ser 165	Met As	sp Pro	Leu Pro	Leu	Pro Leu Ala
	Thr Met						Tyr Thr Ser 168
	Ser Ile 168	Ala Ile	Arg Hi		Glu Arg	Arg	Ile Arg Asn 1695
Asn Arg Ser						Val	Asp Ser Glu 1710
Gly Gly Gly	Asn Val	Thr Ser			Lys Arg	Ser 1725	Thr Arg Asp
		Ser Ser 173	Gln Va	al Glu	Glu Ser 174	Thr	Ser Arg Asp
	Glu Ile						Glu Ala Gln 176
Ala Asp Cys	Glu Leu 176	Thr Asp	Val II		Glu Glu	Glu	
Glu Asp Asp						His	
Arg Arg Lys		Arg Gln			Pro Ser	His	

PCT/US99/29996 WO 00/40614

-29-

60

240 300

360 420

480 489

```
1800
                                              1805
       1795
Thr Asp Leu Ser Glu Gly Glu Glu Val Asp Pro Leu Asp Val Leu Lys
  1810 1815
                                       1820
Met Lys Glu Leu Pro Ile Ile His Gln Ile Leu Asn Glu Glu Gln
     1830
                                    1835
Ala Gly Ala Pro His Ser Thr Pro Val Ile Ala Ser Pro Ser Ser Ser
                        1850
            1845
Arg Ala Asp Leu Thr Ser Gln Lys Cys Ser Asp Val
                              1865
           1860
     <210> 16
     <211> 489
     <212> DNA
     <213> Mus Musculus
ccctgaaaga ctcgacttct gctgctagcg ctggagctga gttagttttg agaaggtttc
ccggggctgt ccttgttcgg tggcccgtgc caccgcctcc ggagacgctt tccgatagat
                                                                   120
ggctgcaggc cgcggaggtg gaggaggagc cgctgccctt ccggagtccg ccccgtgagg
agaatgtccc agaaatcctg gatagagagc actttgacca agagggagtg tgtatatatt
ataccaaget ccaaagacce teacagatgt ettecaggat gteagatttg teageaactt
gtcagatgtt tctgtggtcg tttggtcaag caacatgcat gctttactgc aagtcttgcc
atqaaatact caqatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct
gtggaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg
ggttctcat
     <210> 17
     <211> 102
     <212> PRT
     <213> Mus Musculus
     <400> 17
Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys
                                  10
               5
Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly
                              25
                                                30
           20
Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val
                          40
       35
Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp
                     55
Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val
                                   75
                  70
Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn
                                  90
               85
Phe Gln Gly Gly Ser His
           100
     <210> 18
     <211> 410
     <212> DNA
     <213> Homo Sapiens
     <220>
     <221> unsure
     <222> (6)...(6)
     <221> unsure
     <222> (58)...(58)
     <221> unsure
```

-30-

```
<222> (89)...(89)
      <221> unsure
      <222> (406) ... (406)
      <400> 18
                                                                        60
geogenggag cetgagegga gggtgtgege ageetegeea gegggggeee egggetgnge
cattgcctca ctgagccagc gcctgcctnc tacctcgccg acagctggaa ccagtgcgac
                                                                       120
ctagtggctc tcacctgctt cctcctgggc gtgggctgcc ggctgacccc gggtttgtac
                                                                       180
cacctgggcc gcactgtcct ctgcatcgac ttcatggttt tcacggtgcg gctgcttcac
                                                                       240
atcttcacqq tcaacaaaca gctggggccc aagatcgtca tcgtgagcaa gatgatgaag
                                                                       300
gacgtgttct tcttcctctt cttcctcggc gtgtggctgg tagctatggg ttgggccacg
                                                                       360
gaggggttcc tgaggccacg ggacagtgac ttcccaagta tcctgncgcc
                                                                       410
      <210> 19
      <211> 131
      <212> PRT
      <213> Homo Sapiens
      <220>
      <221> UNSURE
      <222> (15)...(15)
      <223> UNKNOWN
      <221> UNSURE
      <222> (25)...(25)
      <223> UNKNOWN
      <221> UNSURE
      <222> (131)...(131)
      <223> UNKNOWN
      <400> 19
Ala Glu Gly Val Arg Ser Leu Ala Ser Gly Gly Pro Gly Leu Xaa His
                 5
. 1
                                    10
Cys Leu Thr Glu Pro Ala Pro Ala Xaa Tyr Leu Ala Asp Ser Trp Asn
                                25
            20
Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly Cys
        35
                            40
                                                45
Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys Ile
                      55
                                            60
Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn
                    70
                                        75
Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp
                85
                                    90
Val Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Met Gly
           100
                              105
                                                    110
Trp Ala Thr Glu Gly Phe Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser
       115
                           120
Ile Leu Xaa
    130
      <210> 20
      <211> 389
      <212> DNA
      <213> Homo Sapiens
      <400> 20
                                                                       60
caaatttttt gttagtacac catctcatcc aaattgcaaa agtcacatgg aaactggaac
                                                                       120
caaagatcaa gaaactgttt gctctaaagc tacagaagga gataatacag aatttggagc
```

-31-

```
atttqtaqqa cacaqaqata qcatqqatit acaqaqqttt aaaqaaacat caaacaaqat
                                                                    180
aaaaatacta tocaataaca atacttotga aaacactttg aaacgagtga gttotottgo
                                                                    240
tggatttact qactqtcaca qaacttccat tcctgttcat tcaaaacgag aaaagatcag
                                                                    300
tagaaqqcca tctaccgaag acactcatga agtagattcc aaagcagctt taataccggt
                                                                    360
                                                                    389
ttgtagattt caactaaaca gatatatat
     <210> 21
     <211> 415
     <212> DNA
     <213> Homo Sapiens
     <400> 21
atttctagtt tttcaaattt gccagtcttt ttgaatagta tctccttctt ttctcatgtt
                                                                     60
ttatatttaa aacttttta tgtccatcat cactttaaac atacttattt tgtcatctat
                                                                    120
aaccaataat tocactatot tatcagaaat caaataccgt ttatgtaagt tgactcccat
                                                                    180
qaqttctaaa ttqccattqt qaqqtcatct tcqqttaqqc tttaatttqt tqcaaagttq
                                                                    240
                                                                    300
tgcagctcag ggtcaggaag agtccctcca gaaaggagga tttgttactg tgaatctctt
tgttaactaa cctctttccc cactgaaata acttttttca ataacatgat tttaacaaca
                                                                    360
                                                                    415
taatctctct atgccagaac agatatatat gaatgtaagt caatattttc ttgag
     <210> 22
     <211> 405
     <212> DNA
     <213> Mus Musculus
     <400> 22
ttattatggc ttatcatgaa aaaccagtcc tgcctcctcc tcttatcatc ctcagccata
                                                                     60
tagtttcact gttttgctgt gtatgcaaaa gaagaaagaa agataagact tccgatgggc
                                                                    120
caaaactttt cttaacagaa gaagatcaaa agaaactcca tgattttgaa gagcagtgtg
                                                                    180
240
gggtcacttt tgaaagagtg gagcagatga gcattcagat taaagaagtt ggagatcgtg
                                                                    300
tcaactacat aaaaagatca ttacagtctt tagattctca aattggtcat ctgcaagatc
                                                                    360
                                                                    405
tctcaqccct aacagtagat acattgaaaa cacttacagc ccaga
     <210> 23
     <211> 5117
     <212> DNA
     <213> Homo Sapiens
     <220>
     <221> unsure
     <222> (2382)...(2382)
     <223> unknown
     <221> unsure
     <222> (4664)...(4664)
     <223> unknown
     <221> unsure
     <222> (4682)...(4682)
     <223> unknown
     <221> unsure
     <222> (4702)...(4702)
     <223> unknown
     <221> unsure
     <222> (5038)...(5039)
     <223> unknown
```

-32-

<221> unsure <222> (5056)...(5056) <223> unknown <221> unsure <222> (5071)...(5072)

<400> 23

WO 00/40614

qatqqcaaca tgqtqaaqaa tcaatggcta aagcattagt tgcctgtaag atctatcgtt 60 caatggcata tgaagcaaag cagagtgacc tggtagatga tacttcagaa gaactaaaac 120 agtattccaa tgattttggt cagttggccg ttgaattatt agaacagtcc ttcagacaag 180 240 atgaaaccat ggctatgaaa ttgctcactt atgaactgaa gaactggagt aattcaacct geettaagtt ageagtttet teaagaetta gaeettttgt ageteacaee tgtacacaaa 300 tgttgttatc tgatatgtgg atgggaaggc tgaatatgag gaaaaattcc tggtacaagg 360 tcatactaaq cattttaqtt ccacctqcca tattqctqtt agagtataaa actaaggctg 420 aaatgtccca tatcccacaa tctcaagatg ctcatcagat gacaatggat gacagcgaaa 480 540 acaactttca gaacataaca gaagagatcc ccatggaagt gtttaaagaa gtacggattt tggatagtaa tgaaggaaag aatgagatgg agatacaaat gaaatcaaaa aagcttccaa 600 ttacgcgaaa gttttatgcc ttttatcatg caccaattgt aaaattctgg tttaacacgt 660 tggcatattt aggatttctg atgctttata catttgtggt tcttgtacaa atggaacagt 720 taccttcagt tcaagaatgg attgttattg cttatatttt tacttatgcc attgagaaag tccgtgagat ctttatgtct gaagctggga aagtaaacca gaagattaaa gtatggttta 780 840 gtgattactt caacatcagt gatacaattg ccataatttc tttcttcatt ggatttggac 900 960 taagatttgg agcaaaatgg aactttgcaa atgcatatga taatcatgtt tttgtggctg gaagattaat ttactgtctt aacataatat tttggtatgt gcgtttgcta gattttctag 1020 1080 ctgtaaatca acaggcagga ccttatgtaa tgatgattgg aaaaatggtg gccaatatgt tctacattgt agtgattatg gctcttgtat tacttagttt tggtgttccc agaaaggcaa 1140 1200 tactttatcc tcatgaagca ccatcttgga ctcttgctaa agatatagtt tttcacccat 1260 actggatgat ttttggtgaa gtttatgcat acgaaattga tgtgtgtgca aatgattctg ttatccctca aatctgtggt cctgggacgt ggttgactcc atttcttcaa gcagtctacc 1320 totttgtaca gtatatoatt atggttaato ttottattgc atttttcaac aatgtgtatt 1380 1440 tacaagtgaa ggcaatttcc aatattgtat ggaagtacca gcgttatcat tttattatgg cttatcatga gaaaccagtt ctgcctcctc cacttatcat tcttagccat atagtttctc 1500 1560 tgttttgctg catatgtaag agaagaaaga aagataagac ttccgatgga ccaaaacttt tcttaacaga agaagatcaa aagaaacttc atgattttga agagcagtgt gttgaaatgt 1620 1680 atttcaatga aaaagatgac aaatttcatt ctgggagtga agagagaatt cgtgtcactt ttgaaagagt ggaacagatg tgcattcaga ttaaagaagt tggagatcgt gtcaactaca 1740 taaaaagatc attacaatca ttagattctc aaattggcca tttgcaagat ctttcagccc 1800 1860 tgacggtaga tacattaaaa acactcactg cccagaaagc gtcggaagct agcaaagttc ataatgaaat cacacgagaa ctgagcattt ccaaacactt ggctcaaaac cttattgatg 1920 1980 atggtcctgt aagaccttct gtatggaaaa agcatggtgt tgtaaataca cttagctcct 2040 ctcttcctca aggtgatctt gaaagtaata atccttttca ttgtaatatt ttaatgaaag atgacaaaga tccccagtgt aatatatttg gtcaagactt acctgcagta ccccagagaa 2100 2160 aagaatttaa ttttccagag gctggttcct cttctggtgc cttattccca agtgctgttt ccctccaga actgcgacag agactacatg gggtagaact cttaaaaaata tttaataaaa 2220 2280 atcaaaaatt aggcagttca tctactagca taccacatct gtcatcccca ccaaccaaat tttttgttag tacaccatct cagccaagtt gcaaaagcca cttggaaact ggaaccaaag 2340 2400 atcaagaaac tgtttgctct aaagctacag aaggagataa tncagaattt ggagcatttg taggacacag agatagcatg gatttacaga ggtttaaaga aacatcaaac aagataaaaa tactatccaa taacaatact tctgaaaaca ctttgaaacg agtgagttct cttgctggat 2460 2520 ttactgactg tcacagaact tccattcctg ttcattcaaa acaagcagaa aaaatcagta 2580 gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta ataccggatt 2640 ggttacaaga tagaccatca aacagagaaa tgccatctga agaaggaaca ttaaatggtc 2700 tcacttctcc atttaaqcca qctatggata caaattacta ttattcagct gtggaaagaa 2760 ataacttgat gaggttatca cagagcattc catttacacc tgtgcctcca agaggggagc 2820 ctqtcacaqt qtatcqtttq qaaqaqqtt cacccaacat actaaataac agcatgtctt 2880 2940 cttggtcaca actaggcctc tgtgccaaaa tagagttttt aagcaaagag gagatgggag gaggtttacg aagagctgtc aaagtacagt gtacctggtc agaacatgat atcctcaaat cagggcatct ttatattatc aaatcttttc ttccagaggt ggttaataca tggtcaagta 3000 3060 tttataaaga agatacagtt ctgcatctct gtctgagaga aattcaacaa cagagagcag 3120

PCT/US99/29996

-33-

```
cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa
                                                                        3180
ggttccttga agttttcctg ctgtattgcc attcagcagg acagtggttt gctgtggaag
                                                                        3240
aatgtatgac tggagaattt agaaaataca acaataataa tggagatgag attattccaa
                                                                        3300
ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag
                                                                        3360
                                                                        3420
gggagttact ggtacttgat ttgcaaggtg ttggtgaaaa tttgactgac ccatctgtga
                                                                        3480
taaaagcaga agaaaagaga tootgtgata tggtttttgg cocagcaaat ctaggagaag
                                                                        3540
atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac
ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttcct caggatgagc
                                                                        3600
                                                                        3660
cttcagattt gaatcttcag cctggaaatt ccaccaaaga atcagaatca gctaattctg
ttcgtctgat gttataatat taatattact gaatcattgg ttttgcctgc acctcacaga
                                                                        3720
aatgttactg tgtcactttt ccctcgggag gaaattgttt ggtaatatag aaaggtgtat gcaagttgaa tttgctgact ccagcacagt taaaaggtca atattcttt gacctgatta
                                                                        3780
                                                                        3840
atcagtcaga aagtccctat aggatagagc tggcagctga gaaattttaa aggtaattga
                                                                        3900
                                                                        3960
taattagtat ttgtaacttt ttaaagggct ctttgtatag cagaggatct catttgactt
tgttttgatg agggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt
                                                                        4020
gtccatgcag attttattgg cagctgtttt attgccattc aactagggaa atgaagaaat
                                                                        4080
cacgcagcct tttggttaaa tggcagtcaa aattttcctc agtgtattta gtgtgttcag
                                                                        4140
                                                                        4200
tgatgatatc actggttccc aactagatgc ttgttggcca cgggaaggga aatgacttgt
                                                                        4260
tctaattcta ggttcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga
cggtatttat gaatcagggt taggctccat atttaaagat agagccagtt tttttttaa
                                                                        4320
atagaaccca aattgtgtaa aaatgttaat tgggtttttt aaacattgtt ttatcaagtc
                                                                        4380
                                                                        4440
actgttaagt agaagaaagc catggtaaac tgatacataa cctaaattat aaaagcagaa
acctaactca ctcgtcaagg gaagttacct tttgaggaaa gttaaagtac tttttccct
                                                                        4500
atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt
                                                                        4560
tatatcaaat cagaatgtaa acatgaactc ttgcatatat ttaaaattgt gttggaacat
                                                                        4620
                                                                        4680
ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnggatta tcattatgtg
anactggcag attgcagtgc ancettatgc caataaaatg taatttaaca gccccagata
                                                                        4740
                                                                        4800
ttgttgaata ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt
ttttcttttt ttttcttttt cttttgtttc cttggtacta attttaattt ttatttggaa
                                                                        4860
                                                                        4920
gggagcagta taaagcttat ttgtatttag tagtgtatct catagataca gacaaggcaa
                                                                        4980
gagatgataa gctgtttaaa tagtgtttaa tattgattgg gggtggggag aaagaaaaag
tgtattactt aaagatacta tatacgtttt gtatatcatt aaatctttaa aagaaatnna
                                                                        5040
                                                                        5100
ataaatttat tgtttncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag
                                                                        5117
qatecetega ggggeee
```

```
<210> 24
<211> 1224
<212> PRT
<213> Homo Sapiens
<220>
<221> UNSURE
<222> (794)...(794)
<223> UNKNOWN
```

<400> 24 Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys 10 Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp 25 20 Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu 40 35 Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala 60 55 Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys 75 70 Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr -34-

Cys Thr Gln Met Leu Leu Ser Asp Met Trp Met Gly Arg Leu Asn Met Arg Lys Asn Ser Trp Tyr Lys Val Ile Leu Ser Ile Leu Val Pro Pro Ala Ile Leu Leu Glu Tyr Lys Thr Lys Ala Glu Met Ser His Ile Pro Gln Ser Gln Asp Ala His Gln Met Thr Met Asp Asp Ser Glu Asn Asn Phe Gln Asn Ile Thr Glu Glu Ile Pro Met Glu Val Phe Lys Glu 165 170 175 Val Arg Ile Leu Asp Ser Asn Glu Gly Lys Asn Glu Met Glu Ile Gln Met Lys Ser Lys Lys Leu Pro Ile Thr Arg Lys Phe Tyr Ala Phe Tyr His Ala Pro Ile Val Lys Phe Trp Phe Asn Thr Leu Ala Tyr Leu Gly Phe Leu Met Leu Tyr Thr Phe Val Val Leu Val Gln Met Glu Gln Leu Pro Ser Val Gln Glu Trp Ile Val Ile Ala Tyr Ile Phe Thr Tyr Ala 245 250 255 Ile Glu Lys Val Arg Glu Ile Phe Met Ser Glu Ala Gly Lys Val Asn Gln Lys Ile Lys Val Trp Phe Ser Asp Tyr Phe Asn Ile Ser Asp Thr Ile Ala Ile Ile Ser Phe Phe Ile Gly Phe Gly Leu Arg Phe Gly Ala Lys Trp Asn Phe Ala Asn Ala Tyr Asp Asn His Val Phe Val Ala Gly Arg Leu Ile Tyr Cys Leu Asn Ile Ile Phe Trp Tyr Val Arg Leu Leu Asp Phe Leu Ala Val Asn Gln Gln Ala Gly Pro Tyr Val Met Met Ile Gly Lys Met Val Ala Asn Met Phe Tyr Ile Val Val Ile Met Ala Leu Val Leu Leu Ser Phe Gly Val Pro Arg Lys Ala Ile Leu Tyr Pro His Glu Ala Pro Ser Trp Thr Leu Ala Lys Asp Ile Val Phe His Pro Tyr Trp Met Ile Phe Gly Glu Val Tyr Ala Tyr Glu Ile Asp Val Cys Ala Asn Asp Ser Val Ile Pro Gln Ile Cys Gly Pro Gly Thr Trp Leu Thr Pro Phe Leu Gln Ala Val Tyr Leu Phe Val Gln Tyr Ile Ile Met Val Asn Leu Leu Ile Ala Phe Phe Asn Asn Val Tyr Leu Gln Val Lys Ala Ile Ser Asn Ile Val Trp Lys Tyr Gln Arg Tyr His Phe Ile Met Ala Tyr His Glu Lys Pro Val Leu Pro Pro Pro Leu Ile Ile Leu Ser His Ile Val Ser Leu Phe Cys Cys Ile Cys Lys Arg Arg Lys Lys Asp Lys Thr Ser Asp Gly Pro Lys Leu Phe Leu Thr Glu Glu Asp Gln Lys Lys Leu His Asp Phe Glu Glu Gln Cys Val Glu Met Tyr Phe Asn Glu Lys Asp Asp Lys Phe His Ser Gly Ser Glu Glu Arg Ile Arg Val Thr Phe Glu Arg Val Glu Gln Met Cys Ile Gln Ile Lys Glu Val Gly Asp Arg

Val	Asn	Tyr	Ile 580	Lys	Arg	Ser	Leu	Gln 585	Ser	Leu	Asp	Ser	Gln 590	Ile	Gly
His	Leu	Gln 595	Asp	Leu	Ser	Ala	Leu 600	Thr	Val	Asp	Thr	Leu 605	Lys	Thr	Leu
Thr			Lys	Ala	Ser	Glu 615		Ser	Lys	Val	His 620		Glu	Ile	Thr
-	610 Glu	Leu	Ser	Ile	Ser		His	Leu	Ala	Gln 635		Leu	Ile	Asp	Asp 640
625 Gly	Pro	Val	Arg		630 Ser	Val	Trp	Lys			Gly	Val	Val		
Leu	Ser	Ser		645 Leu	Pro	Gln	Gly		650 Leu	Glu	Ser	Asn		655 Pro	Phe
His	Cys		660 Ile	Leu	Met	Lys		665 Asp	Lys	Asp	Pro		670 Cys	Asn	Ile
Phe		675 Gln	Asp	Leu	Pro		680 Val	Pro	Gln	Arg		685 Glu	Phe	Asn	Phe
	690 Glu	Ala	Gly	Ser	Ser	695 Ser	Gly	Ala	Leu		700 Pro	Ser	Ala	Val	
705 Pro	Pro	Glu	Leu	Arg	710 Gln	Arg	Leu	His		715 Val	Glu	Leu	Leu		720 Ile
Phe	Asn	Lys	Asn	725 Gln	Lys	Leu	Gly		730 Ser	Ser	Thr	Ser		735 Pro	His
Leu	Ser	Ser	740 Pro	Pro	Thr	Lys	Phe	745 Phe	Val	Ser	Thr		750 Ser	Gln	Pro
Ser	Cys	755 Lys	Ser	His	Leu		760 Thr	Gly	Thr	Lys		765 Gln	Glu	Thr	Val
Cys	770 Ser	Lys	Ala	Thr	Glu	775 Gly	Asp	Asn	Xaa	Glu	780 Phe	Gly	Ala	Phe	Val
785					790					795					800
Gly	His	Arg	Asp	Ser 805	Met	Asp	Leu	Gln	Arg 810	Phe	Lys	Glu	Thr	Ser 815	Asn
Lys	Ile	Lys	Ile 820	Leu	Ser	Asn	Asn	Asn 825	Thr	Ser	Glu	Asn	Thr 830	Leu	Lys
Arg	Val	Ser 835	Ser	Leu	Ala	Gly	Phe 840	Thr	Asp	Суѕ	His	Arg 845	Thr	Ser	Ile
Pro	Val 850	His	Ser	Lys	Gln	Ala 855	Glu	Lys	Ile	Ser	Arg 860	Arg	Pro	Ser	Thr
Glu 865	Asp	Thr	His	Glu	Val 870	Asp	Ser	Lys	Ala	Ala 875	Leu	Ile	Pro	Asp	Trp 880
Leu	Gln	Asp	Arg	Pro 885	Ser	Asn	Arg	Glu	Met 890	Pro	Ser	Glu	Glu	Gly 895	Thr
Leu	Asn	Gly	Leu 900	Thr	Ser	Pro	Phe	Lys 905	Pro	Ala	Met	Asp	Thr 910	Asn	Tyr
Tyr	Tyr	Ser 915		Val	Glu	Arg	Asn 920	Asn	Leu	Met	Arg	Leu 925	Ser	Gln	Ser
Ile	Pro 930	Phe	Thr	Pro	Val	Pro 935	Pro	Arg	Gly	Glu	Pro 940	Val	Thr	Val	Tyr
Arg 945	Leu	Glu	Glu	Ser	Ser 950	Pro	Asn	Ile	Leu	Asn 955	Asn	Ser	Met	Ser	Ser 960
Trp	Ser	Gln	Leu	Gly 965	Leu	Cys	Ala	Lys	Ile 970	Glu	Phe	Leu	Ser	Lys 975	Glu
Glu	Met	Gly	Gly 980	Gly	Leu	Arg	Arg	Ala 985	Val	Lys	Val	Gln	Cys 990	Thr	Trp
Ser	Glu	His 995		Ile	Leu	Lys	Ser 1000		His	Leu	Tyr	Ile 1005		Lys	Ser
Phe	Leu 1010	Pro	Glu	Val	Val	Asn 1015	Thr		Ser	Ser	Ile 1020		Lys	Glu	Asp
Thr 1025	Val		His	Leu	Cys 1030	Leu		Glu	Ile	Gln 1035	Gln		Arg	Ala	Ala 104
		Leu	Thr	Phe 1045	Ala		Asn	Gln	Met 1050	Lys		Lys	Ser	Ile 1055	Pro

-36-Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala 1070 1060 1065 Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys 1075 1080 1085 Tyr Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu 1095 1100 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly 1110 1115 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp 1125 1130 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe 1140 1145 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys 1160 1165 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys 1170 1175 1180 Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro 1190 1195 Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser 1205 1210 Ala Asn Ser Val Arg Leu Met Leu 1220

<210> 25 <211> 2180 <212> DNA

<213> Homo Sapiens

<400> 25

tcgaggccaa gaattcggca cgagggcctc gggcaggccc cctggagcga cctgcttctt 60 120 tgggcactgt tgctgaacag ggcacagatg gccatgtact tctgggagat gggttccaat 180 granttteet cagetettgg ggeetgtttg etgeteeggg tgatggeaeg eetggageet 240 gacgctgagg aggcagcacg gaggaaagac ctggcgttca agtttgaggg gatgggcgtt 300 gacctctttg gcgagtgcta tcgcagcagt gaggtgaggg ctgcccgcct cctcctccgt 360 cgctgcccgc tctgggggga tgccacttgc ctccagctgg ccatgcaagc tgacgcccgt 420 gccttctttg cccaggatgg ggtacagtct ctgctgacac agaagtggtg gggagatatg gccagcacta cacccatctg ggccctggtt ctcgccttct tttgccctcc actcatctac 480 accegeetea teacetteag gaaateagaa gaggageeea eaegggagga getagagttt 540 600 gacatggata gtgtcattaa tggggaaggg cctgtcggga cggcggaccc agccgagaag acgccgctgg gggtcccgcg ccagtcgggc cgtccgggtt gctgcggggg ccgctgcggg 660 720 gggcgccqgt gcctacqccg ctggttccac ttctggggcg cgccggtgac catcttcatg 780 ggcaacgtgg tcagctacct gctgttcctg ctgcttttct cgcgggtgct gctcgtggat 840 ttccagccgg cgccgcccgg ctccctggag ctgctgctct atttctgggc tttcacgctg ctgtgcgagg aactgcgcca gggcctgagc ggaggcgggg gcagcctcgc cagcgggggc 900 960 cccgggcctg gccatgcctc actgagccag cgcctgcgcc tctacctcgc cgacagctgg aaccagtgcg acctagtggc teteacetgc tteeteetgg gegtgggetg eeggetgace 1020 1080 coggqtttqt accacctqqq ccqcactqtc ctctqcatcq acttcatqqt tttcacqqtq 1140 cggctgcttc acatcttcac ggtcaacaaa cagctggggc ccaagatcgt catcgtgagc 1200 aagatgatga aqqacqtqtt cttcttcctc ttcttcctcg gcgtgtggct ggtagcctat 1260 ggcgtggcca cggaggggct cctgaggcca cgggacagtg acttcccaag tatcctgcgc cgcgtcttct accgtcccta cctgcagatc ttcgggcaga ttccccagga ggacatggac 1320 gtggccctca tggagcacag caactgctcg tcggagcccg gcttctgggc acaccctcct 1380 1440 ggggcccagg cgggcacctg cgtctcccag tatgccaact ggctggtggt gctgctcctc gtcatcttcc tgctcgtggc caacatcctg ctggtcaact tgctcattgc catgttcagt 1500 1560 tacacatteg gcaaagtaca gggcaacage gatetetact ggaaggegca gegttacege 1620 ctcatccggg aattccactc tcggcccgcg ctggccccgc cctttatcgt catctcccac ttgcgcctcc tgctcaggca attgtgcagg cgacccsgga gcccccagcc gtcctccccg 1680 1740 gccctcgagc atttccgggt ttacctttct aaggaagccg agcggaagct gctaacgtgg 1800 gaatcggtgc ataaggagaa ctttctgctg gcacgcgcta gggacaagcg ggagagcgac tecqaqmqtc tqaaqcqcac qteccaqaaq gtggacttgg cactgaaaca gctgggacac 1860

PCT/US99/29996 WO 00/40614 -37-

atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg 1920 geoccegcae etggtggeet tgteettgag gtgageecca tgteeatetg ggeeactgte 1980 aggaccacct ttgggagtgt catccttaca aaccacagca tgcccggctc ctcccagaac 2040 cagteceage etgggaggat caaggeetgg atccerggee gttatecate tggaggetge 2100 2160 agggtccttg gggtaacagg gaccacagac ccctcaccac tcacagattc ctcacactgg 2180 ggaaataaag ccatttcaga

<210> 26 <211> 725 <212> PRT <213> Homo Sapiens <220> <221> UNSURE <222> (553)...(553) <223> UNKNOWN

<221> UNSURE

<222> (603)...(603)

<223> UNKNOWN

195

<400> 26 Ser Arg Pro Arg Ile Arg His Glu Gly Leu Gly Gln Ala Pro Trp Ser 10 15 Asp Leu Leu Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met 25 Tyr Phe Trp Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala 40 35 Cys Leu Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu 55 60 Ala Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val 70 75 Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala Arg 90 85 Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys Leu Gln 100 105 110 Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln Asp Gly Val 120 Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met Ala Ser Thr Thr 135 140 130 Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys Pro Pro Leu Ile Tyr 155 150 Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu Glu Glu Pro Thr Arg Glu 175 165 170 Glu Leu Glu Phe Asp Met Asp Ser Val Ile Asn Gly Glu Gly Pro Val 180 185 Gly Thr Ala Asp Pro Ala Glu Lys Thr Pro Leu Gly Val Pro Arg Gln 200

Ser Gly Arg Pro Gly Cys Cys Gly Gly Arg Cys Gly Gly Arg Arg Cys 215 220 210 Leu Arg Arg Trp Phe His Phe Trp Gly Ala Pro Val Thr Ile Phe Met 235 230 Gly Asn Val Val Ser Tyr Leu Leu Phe Leu Leu Phe Ser Arg Val 250 245 Leu Leu Val Asp Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu 265 270 260 Leu Tyr Phe Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly 280 275 285 Leu Ser Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly 295 300

PCT/US99/29996 WO 00/40614

****	00/40	· · · · · · · · · · · · · · · · · · ·													
									-38-						
305			Leu		310					315					320
Asn	Gln	Cys	Asp	Leu 325	Val	Ala	Leu	Thr	Cys 330	Phe	Leu	Leu	Gly	Val 335	Gly
Cys	Arg	Leu	Thr 340	Pro	Gly	Leu	Tyr	His 345	Leu	Gly	Arg	Thr	Val 350	Leu	Суѕ
Ile	Asp	Phe 355	Met	Val	Phe	Thr	Val 360	Arg	Leu	Leu	His	Ile 365	Phe	Thr	Val
Asn	Lys 370	Gln	Leu	Gly	Pro	Lys 375	Ile	Val	Ile	Val	Ser 380	Lys	Met	Met	Lys
385			Phe		390					395					400
			Thr	405					410					415	
			Arg 420					425					430		
		435	Gln				440					445			
	450		Glu			455					460				
465			Val		470					475					480
			Leu	485					490					495	
			Ser 500					505					510		
		515	Ala				520					525			
	530		Ala			535					540				
545			Leu		550					555					560
			His	565					570					575	
			Trp 580					585					590		
	-	595	Lys				600					605			
	610		Asp Leu			615					620				
625			Pro		630					635					640
			Val	645					650					655	
			660 Gly					665					670		
		675	Pro				680					685			
	690		Thr			695					700				
705			Ala		710	FIO	261	110	neu	715	лор	501			720
ату	Moll	ъys	VIG	725											

<210> 27 <211> 7419 <212> DNA <213> Homo Sapiens

PCT/US99/29996 WO 00/40614 -39-

<400> 27 60 cggggaccga tccagcctcc ggactctagc ctaggctttt gcaaaaagct atttaggtga cactatagaa ggtacgcctg caggtaccgg tccggaattc ccgggtcgac ccacgcgtcc 120 180 qcaqcccqt cgccggcgga ggcgggcgcg ggcgcgtncc ctgtggccag tcacccggag 240 gagttggtcg cacaattatg aaagactcgg cttctgctgc tagcgccgga gctgagttag ttctgagaag gtttccctgg gcgttccttg tccggcggcc tctgctgccg cctccggaga 300 360 cgcttcccga tagatggcta caggccgcgg aggaggagga ggtggagttg ctgcccttcc ggagtccgcc ccgtgaggag aatgtcccag aaatcctgga tagaaagcac tttgaccaag 420 480 agggaatgtg tatatattat accaagttcc aaggaccctc acagatgcct tccaggatgt caaatttgtc agcaactcgt caggtgtttt tgtggtcgct tggtcaagca acatgcttgt 540 600 tttactgcaa gtcttgccat gaaatactca gatgtgaaat tgggtgacca ttttaatcag 660 720 gtcataaatt ttcaaggggg ttctcattcc tacagagcta agtatgtgag gctatcatat gacaccaaac ctgaagtcat totgcaactt ctgottaaag aatggcaaat ggagttacco 780 840 aaacttgtta tetetgtaca tgggggcatg cagaaatttg agetteacce acgaatcaag cagttgcttg gaaaaggtct tattaaagct gcagttacaa ctggagcctg gattttaact 900 960 qqaqqaqtaa acacaggtgt ggcaaaacat gttggagatg ccctcaaaga acatgcttcc 1020 agatcatctc gaaagatttg cactatcgga atagctccat ggggagtgat tgaaaacaga 1080 aatgatettg ttgggagaga tgtggttget eettateaaa eettattgaa eeceetgage aaattgaatg ttttgaataa tctgcattcc catttcatat tggtggatga tggcactgtt 1140 ggaaagtatg gggcggaagt cagactgaga agagaacttg aaaaaactat taatcagcaa 1200 agaattcatg ctaggattgg ccagggtgtc cctgtggtgg cacttatatt tgagggtggg 1260 1320 ccaaatqtta tcctcacaqt tcttgaatac cttcaggaaa gcccccctgt tccagtagtt 1380 qtqtqtqaag gaacaggcag agctgcagat ctgctagcgt atattcataa acaaacagaa gaaggagga atcttcctga tgcagcagag cccgatatta tttccactat caaaaaaaca 1440 tttaactttg gccagaatga agcacttcat ttatttcaaa cactgatgga gtgcatgaaa 1500 agaaaggagc ttatcactgt tttccatatt gggtcagatg aacatcaaga tatagatgta 1560 1620 gcaatactta ctgcactgct aaaaggtact aatgcatctg catttgacca gcttatcctt acattggcat gggatagagt tgacattgcc aaaaatcatg tatttgttta tggacagcag 1680 1740 tggctggttg gatccttgga acaagctatg cttgatgctc ttgtaatgga tagagttgca 1800 tttgtaaaac ttcttattga aaatggagta agcatgcata aattccttac cattccgaga 1860 ctggaagaac tttacaacac taaacaaggt ccaactaatc caatgctgtt tcatcttgtt cgagacgtca aacagggaaa tcttcctcca ggatataaga tcactctgat tgatatagga 1920 1980 cttgttattg aatatctcat gggaggaacc tacagatgca cctatactag gaaacgtttt 2040 cgattaatat ataatagtct tggtggaaat aatcggaggt ctggccgaaa tacctccagc 2100 agcactcctc agttgcgaaa gagtcatgaa tcttttggca atagggcaga taaaaaggaa 2160 aaaatgaggc ataaccattt cattaagaca gcacagccct tccgaccaaa gattgataca 2220 gttatggaag aaggaaagaa gaaaagaacc aaagatgaaa ttgtagacat tgatgatcca 2280 gaaaccaagc gctttcctta tccacttaat gaacttttaa tttgggcttg ccttatgaag 2340 aggcaggtca tggcccgttt tttatggcaa catggtgaag aatcaatggc taaagcatta gttgcctgta agatctatcg ttcaatggca tatgaagcaa agcagagtga cctggtagat 2400 2460 gatacttcag aagaactaaa acagtattcc aatgattttg gtcagttggc cgttgaatta ttagaacagt ccttcagaca agatgaaacc atggctatga aattgctcac ttatgaactg 2520 2580 aagaactgga gtaattcaac ctgccttaag ttagcagttt cttcaagact tagacctttt gtagctcaca cctgtacaca aatgttgtta tctgatatgt ggatgggaag gctgaatatg 2640 2700 aggaaaaatt cctggtacaa ggtcatacta agcattttag ttccacctgc catattgctg ttagagtata aaactaagge tgaaatgtee catateeeac aateteaaga tgeteateag 2760 2820 atgacaatgg atgacagcga aaacaacttt cagaacataa cagaagagat ccccatggaa gtgtttaaag aagtacggat tttggatagt aatgaaggaa agaatgagat ggagatacaa 2880 2940 atgaaatcaa aaaagcttcc aattacgcga aagttttatg ccttttatca tgcaccaatt 3000 gtaaaattct ggtttaacac gttggcatat ttaggatttc tgatgcttta tacatttgtg 3060 qttcttgtac aaatggaaca gttaccttca gttcaagaat ggattgttat tgcttatatt tttacttatg ccattgagaa agtccgtgag atctttatgt ctgaagctgg gaaagtaaac 3120 3180 cagaagatta aagtatggtt tagtgattac ttcaacatca gtgatacaat tgccataatt 3240 tetttettea ttggatttgg actaagattt ggagcaaaat ggaaetttge aaatgeatat gataatcatg tttttgtggc tggaagatta atttactgtc ttaacataat attttggtat 3300 3360 gtgcgtttgc tagattttct agctgtaaat caacaggcag gaccttatgt aatgatgatt ggaaaaatgg tggccaatat gttctacatt gtagtgatta tggctcttgt attacttagt 3420 3480 tttggtgttc ccagaaaggc aatactttat cctcatgaag caccatcttg gactcttgct aaagatatag tttttcaccc atactggatg atttttggtg aagtttatgc atacgaaatt 3540

-40-

•						2600
gatgtgtgtg	caaatgattc	tgttatccct	caaatctgtg	gtcctgggac	gtggttgact	3600
ccatttcttc	aagcagtcta	cctctttgta	cagtatatca	ttatggttaa	tcttcttatt	3660
gcatttttca	acaatgtgta	tttacaagtg	aacqcaattt	ccaatattqt	atggaagtac	3720
	attttattat					3780
	atatagtttc					3840
attettagee	acacageeee	tergereege	cgcacacgca	agagaagaaa	tastastttt	3900
acttccgatg	gaccaaaact	tttcttaaca	gaagaagatc	aaaagaaact	tcatgattt	
	gtgttgaaat					3960
	ttcgtgtcac					4020
gttggagatc	gtgtcaacta	cataaaaaga	tcattacaat	cattagattc	tcaaattggc	4080
	atctttcagc					4140
	ctagcaaagt					4200
ttaaataaaa	accttattga	tastaataat	gtaagacett	ctatataaaa	aaaacataat	4260
rtggccaaa	accttattya	cyacyycccc	graagacere	ttannatan	taateettt	4320
	cacttagctc					
	ttttaatgaa					4380
	taccccagag					4440
gccttattcc	caagtgctgt	ttcccctcca	gaactgcgac	agagactaca	tggggtagaa	4500
ctcttaaaaa	tatttaataa	aaatcaaaaa	ttaggcagtt	catctactag	cataccacat	4560
ctgtcatccc	caccaaccaa	attttttgtt	agtacaccat	ctcagccaag	ttgcaaaagc	4620
cacttggaaa	ctggaaccaa	acatcaacaa	actotttoct	ctaaagctac	agaaggagat	4680
cacciggaaa	ttggagcatt	tatagaa	accecteges	tagatttaga	agaaggagaa	4740
aatacagaat	Liggageaci	tytayyatat	agagatagta	tygatttata	gaggeetaaa	4800
	acaagataaa					
	ctcttgctgg					4860
aaacaagcag	aaaaaatcag	tagaaggcca	tctaccgaag	acactcatga	agtagattcc	4920
aaagcagctt	taataccgga	ttggttacaa	gatagaccat	caaacagaga	aatgccatct	4980
	cattaaatgg					5040
	ctgtggaaag					5100
cetatacete	caagagggga	acctatcaca	atatatcatt	tagaagagag	ttcacccaac	5160
atagtgaata	acagcatgtc	ttetteetee	gagatagga	tetataccaa	aatagagttt	5220
atactadata	acagcatytt	Lectiggica	caactaggtt	teracetees	aacagageee	5280
ttaagcaaag	aggagatggg	aggaggttta	cgaagagetg	tcaaagtaca	gracerag	
tcagaacatg	atatcctcaa	atcagggcat	ctttatatta	tcaaatcttt	tcttccagag	5340
gtggttaata	catggtcaag	tatttataaa	gaagatacag	ttctgcatct	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacgtttg	cctttaatca	aatgaaaccc	5460
aaatccatac	catattctcc	aaggttcctt	gaagttttcc	tgctgtattg	ccattcagca	5520
ggacagtggt	ttgctgtgga	agaatgtatg	actggagaat	ttagaaaata	caacaataat	5580
aatggagatg	agattattcc	aactaatact	ctggaagaga	tcatgctagc	ctttagccac	5640
togacttaco	aatatacaag	aggggagtta	ctggtacttg	atttgcaagg	tattaataaa	5700
	acccatctgt					5760
	atctaggaga					5820
ggcccagcaa	acctaggaga	agatgeaact	andacccca	atgattatag	gastastasa	5880
torraction	gaaagcttaa	acticcagat	ttgaagagga	acquecacac	ttaasaassa	5940
attatatttc	ctcaggatga	geetteagat	tigaatette	agectygaaa	LLCCaccaaa	
gaatcagaat	caactaattc	tgttcgtctg	atgttataat	attaatatta	ctgaatcatt	6000
ggttttgcct	gcacctcaca	gaaatgttac	tgtgtcactt	ttccctcggg	aggaaattgt	6060
ttggtaatat	agaaaggtgt	atgcaagttg	aatttgctga	ctccagcaca	gttaaaaggt	6120
caatattctt	ttgacctgat	taatcagtca	gaaagtccct	ataggataga	gctggcagct	6180
gagaaatttt	aaaggtaatt	gataattagt	atttgtaact	ttttaaaggg	ctctttgtat	6240
agragaggat	ctcatttgac	tttattttaa	tgagggtgat	accetetett	atgtggtaca	6300
ataccattaa	ccaaaggtag	atatacetac	agattttatt	accacetatt	ttattqccat	6360
tesestes	ccaaaggtag	gtgtccatge	attttaatta	antagoogta	anantttee	6420
tcaactaggg	aaatgaagaa	arcacgcage	Cccctygcca	aacggcagcc	aaaaccccc	
tcagtgtatt	tagtgtgttc	agtgatgata	tcactggttc	ccaactagat	gettgttgge	6480
cacgggaagg	gaaatgactt	gttctaattc	taggttcaca	gaggtatgag	aagcctgaac	6540
	tttcaagagg					6600
atagagccag	tttttttt	aaatagaacc	caaattgtgt	aaaaatgtta	attgggtttt	6660
ttaaacattg	ttttatcaag	tcactqttaa	gtagaagaaa	gccatggtaa	actgatacat	6720
	ataaaagcag					6780
aanttaaant	acttttttcc	ctatctctatat	ctataggaag	aacccagaac	ttacaaactt	6840
atanaast	tttattgatt	attatatasa	atcacaatct	2220249220	tettacatat	6900
	gtgttggaac					6960
	tatcattatg					7020
	cagccccaga					7080
ctaagtttta	tgctttaatt	ttttttcttt	ttttttcttt	ttcttttgtt	tccttggtac	7140

PCT/US99/29996 WO 00/40614

7200

7260 7320

7380

7419

taattttaat ttttatttgg aagggagcag tataaagctt atttgtattt agtagtgtat ctcatagata cagacaaggc aagagatgat aagctgttta aatagtgktt aatattgatt qqqqqtqqqq agaaagaaaa agtgtattac ttaaagatac tatatacskt ttktatatca ttaaatcttt aaaagaaatn naataaattt attgtttnca aaaaaaaaac ccnntaaaaa

> <210> 28 <211> 1865 <212> PRT <213> Homo Sapiens

aaaaagggcg gccctctag aggatccctc gaggggccc

<400> 28

Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys 10 Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly 20 25 30 Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val 40 45 Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp 55 Val Lys Leu Gly Asp His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val 75 Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn 90 85 Phe Gln Gly Gly Ser His Ser Tyr Arg Ala Lys Tyr Val Arg Leu Ser 105 Tyr Asp Thr Lys Pro Glu Val Ile Leu Gln Leu Leu Leu Lys Glu Trp 120 115 Gln Met Glu Leu Pro Lys Leu Val Ile Ser Val His Gly Gly Met Gln 135 140 Lys Phe Glu Leu His Pro Arg Ile Lys Gln Leu Leu Gly Lys Gly Leu 150 155 Ile Lys Ala Ala Val Thr Thr Gly Ala Trp Ile Leu Thr Gly Gly Val 170 165 Asn Thr Gly Val Ala Lys His Val Gly Asp Ala Leu Lys Glu His Ala 185 180 Ser Arg Ser Ser Arg Lys Ile Cys Thr Ile Gly Ile Ala Pro Trp Gly 200 Val Ile Glu Asn Arg Asn Asp Leu Val Gly Arg Asp Val Val Ala Pro 215 210 Tyr Gln Thr Leu Leu Asn Pro Leu Ser Lys Leu Asn Val Leu Asn Asn 235 230 Leu His Ser His Phe Ile Leu Val Asp Asp Gly Thr Val Gly Lys Tyr 250 245 Gly Ala Glu Val Arg Leu Arg Arg Glu Leu Glu Lys Thr Ile Asn Gln 265 260 Gln Arg Ile His Ala Arg Ile Gly Gln Gly Val Pro Val Val Ala Leu 280 285 275 Ile Phe Glu Gly Gly Pro Asn Val Ile Leu Thr Val Leu Glu Tyr Leu 300 295 Gln Glu Ser Pro Pro Val Pro Val Val Val Cys Glu Gly Thr Gly Arg 310 315 Ala Ala Asp Leu Leu Ala Tyr Ile His Lys Gln Thr Glu Glu Gly Gly 330 325 335 Asn Leu Pro Asp Ala Ala Glu Pro Asp Ile Ile Ser Thr Ile Lys Lys 345 350 Thr Phe Asn Phe Gly Gln Asn Glu Ala Leu His Leu Phe Gln Thr Leu 360 365

-42-Met Glu Cys Met Lys Arg Lys Glu Leu Ile Thr Val Phe His Ile Gly Ser Asp Glu His Gln Asp Ile Asp Val Ala Ile Leu Thr Ala Leu Leu Lys Gly Thr Asn Ala Ser Ala Phe Asp Gln Leu Ile Leu Thr Leu Ala Trp Asp Arg Val Asp Ile Ala Lys Asn His Val Phe Val Tyr Gly Gln Gln Trp Leu Val Gly Ser Leu Glu Gln Ala Met Leu Asp Ala Leu Val Met Asp Arg Val Ala Phe Val Lys Leu Leu Ile Glu Asn Gly Val Ser Met His Lys Phe Leu Thr Ile Pro Arg Leu Glu Glu Leu Tyr Asn Thr Lys Gln Gly Pro Thr Asn Pro Met Leu Phe His Leu Val Arg Asp Val Lys Gln Gly Asn Leu Pro Pro Gly Tyr Lys Ile Thr Leu Ile Asp Ile Gly Leu Val Ile Glu Tyr Leu Met Gly Gly Thr Tyr Arg Cys Thr Tyr Thr Arg Lys Arg Phe Arg Leu Ile Tyr Asn Ser Leu Gly Gly Asn Asn Arg Arg Ser Gly Arg Asn Thr Ser Ser Ser Thr Pro Gln Leu Arg Lys Ser His Glu Ser Phe Gly Asn Arg Ala Asp Lys Lys Glu Lys Met Arg His Asn His Phe Ile Lys Thr Ala Gln Pro Phe Arg Pro Lys Ile Asp Thr Val Met Glu Glu Gly Lys Lys Lys Arg Thr Lys Asp Glu Ile Val Asp Ile Asp Asp Pro Glu Thr Lys Arg Phe Pro Tyr Pro Leu Asn Glu Leu Leu Ile Trp Ala Cys Leu Met Lys Arg Gln Val Met Ala Arg Phe Leu Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys
645 650 655 Lys Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr Cys Thr Gln Met Leu Leu Ser Asp Met Trp Met Gly Arg Leu Asn Met Arg Lys Asn Ser Trp Tyr Lys Val Ile Leu Ser Ile Leu Val Pro Pro Ala Ile Leu Leu Glu Tyr Lys Thr Lys Ala Glu Met Ser His Ile Pro Gln Ser Gln Asp Ala His Gln Met Thr Met Asp Asp Ser Glu Asn Asn Phe Gln Asn Ile Thr Glu Glu Ile Pro Met Glu Val Phe Lys Glu Val Arg Ile Leu Asp Ser Asn Glu Gly Lys Asn Glu Met Glu Ile Gln Met Lys Ser Lys Lys Leu Pro Ile Thr Arg Lys Phe Tyr Ala Phe

-	850					855	Phe				860				
865	Phe	Leu			870					875					880
Leu		Ser		885					890					895	
		Glu	900					905					910		
		Lys 915					920					925			
	930	Ala				935					940				
945		Trp			950					955					960
		Leu		965					970					975	
		Phe	980					985					990		
		Lys 995					1000)				1005)		
	1010	Leu)				1015	5				1020)			
1025	5	Ala			1030)				1035	5				1040
		Met		1045	5				1050)				1055	Ó
		Asp	1060)				1065	5				1070)	
		Phe 1079	5				1080)				1085	5		
	1090	Leu)				1095	5				1100)			
1105	5	Ser			1110)				1115	5				1120
		His		1125	5				1130)				1135	5
		Val	1140)				1145	5				1150)	
		Ser 1155	5				1160)				1165	5		
	1170	His O				1175	5				1180)			
1189	5	Asp			1190)				1195	5				1200
		Arg		1205	5				1210	0				121	5
		Asn	1220)				1225	5				1230)	
		Leu 123	5				1240)				1245	5		
	1250	Ala O				1255	5				1260)			
126	5	Glu			1270)				127	5				1280
•	_	Pro		1285	5				129	0				129	5
		Ser	1300)				1305	5				1310)	
Phe	His	Cys 131		Ile	Leu	Met	Lys 1320		Asp	Lys	Asp	Pro 132	Gln 5	Cys	Asn

Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn 1335 1340 Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val 1350 1355 1345 Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys 1365 1370 1375 Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro 1380 1385 1390 His Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln 1395 1400 1405 Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr 1410 1415 1420 Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe 1430 1435 Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser 1445 1450 1455 Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu 1460 1465 1470 Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser 1480 1485 1475 Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser 1495 1500 Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp 1510 1515 Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly 1525 1530 1535 Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn 1550 1545 1540 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln 1555 1560 1565 Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val 1575 1580 1570 Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser 1590 1595 1600 Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys 1605 1610 1615 Glu Glu Met Gly Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr 1620 1625 1630 Trp Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys 1635 1640 1645 Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu 1650 1655 1660 Asp Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Arg Ala 1675 1665 1670 Ala Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile 1685 1690 Pro Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser 1710 1700 1705 Ala Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg 1725 1715 1720 Lys Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu 1730 1735 1740 Glu Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg 1745 1750 1755 Gly Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr 1765 1770 1775 Asp Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val 1785 1790 1780 Phe Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala 1800 1805

-45-

<210> 29 <211> 4061 <212> DNA <213> Homo Sapiens

<400> 29

qqtctqqaaq caqaqccqqc ggagggagcg ccggggccct gggctgcagg aggttgcggc 60 ggccgcggca gcatggtggt gccggagaag gagcagagct ggatccccaa gatcttcaag 120 aaqaaqacct gcacqacqtt cataqttgac tccacagatc cgggagggac cttgtgccag 180 240 tqtqqqcqcc cccqqaccqc ccaccccqca gtqqccatqq aggatqcctt cqgqqcaqcc qtqqtqaccq tqtqqqacaq cqatqcacac accacggaga agcccaccga tqcctacgga 300 360 gagctggact tcacgggggc cggccgcaag cacagcaatt tcctccggct ctctgaccga acggatccag ctgcagttta tagtctggtc acacgcacat ggggcttccg tgccccgaac ctggtggtgt cagtgctggg gggatcgggg ggccccgtcc tccagacctg gctgcaggac 420 480 ctgctgcgtc gtgggctggt gcgggctgcc cagagcacag gagcctggat tgtcactggg 540 600 ggtctgcaca cgggcatcgg ccggcatgtt ggtgtggctg tacgggacca tcagatggcc agcactgggg gcaccaaggt ggtggccatg ggtgtggccc cctggggtgt ggtccggaat agagacacc tcatcaacc caagggctcg ttccctgcga ggtaccggtg gcgcggtgac 660 720 ccggaggacg gggtccagtt tcccctggac tacaactact cggccttctt cctggtggac 780 gacggcacac acggctgcct ggggggcgag aaccgcttcc gcttgcgcct ggagtcctac 840 atctcacagc agaagacggg cgtgggaggg actggaattg acatccctgt cctgctcctc 900 ctgattgatg gtgatgagaa gatgttgacg cgaatagaga acgccaccca ggctcagctc 960 1020 ccatqtctcc tcqtqqctqq ctcaggggga gctqcggact gcctggcgga gaccctggaa 1080 gacactctgg ccccagggag tgggggagcc aggcaaggcg aagcccgaga tcgaatcagg cgtttctttc ccaaagggga ccttgaggtc ctgcaggccc aggtggagag gattatgacc cggaaggagc tcctgacagt ctattcttct gaggatgggt ctgaggaatt cgagaccata 1140 1200 qttttqaaqq cccttqtqaa qqcctqtqqq agctcggagg cctcagccta cctggatgag 1260 1320 ctgcgtttgg ctgtggcttg gaaccgcgtg gacattgccc agagtgaact ctttcggggg gacatccaat ggcggtcctt ccatctcgaa gcttccctca tggacgccct gctgaatgac 1380 cggcctgagt tcgtgcgctt gctcatttcc cacggcctca gcctgggcca cttcctgacc 1440 1500 cegatgegee tggeceaact ctacagegeg gegeeeteea actegeteat cegeaacett 1560 ttggaccagg cgtcccacag cgcaggcacc aaagccccag ccctaaaagg gggagctgcg gageteegge eccetgaegt ggggeatgtg etgaggatge tgetggggaa gatgtgegeg 1620 1680 ccgaggtacc cctccggggg cgcctgggac cctcacccag gccagggctt cggggagagc atgtatctgc tctcggacaa ggccacctcg ccgctctcgc tggatgctgg cctcgggcag 1740 1800 qccccctqqa qcqacctqct tctttqgqca ctgttgctga acagggcaca gatggccatg tacttetggg agatgggtte caatgeagtt teeteagete ttggggeetg tttgetgete 1860 cgggtgatgg cacgcctgga gcctgacgct gaggaggcag cacggaggaa agacctggcg ttcaagtttg aggggatggg cgttgacctc tttggcgagt gctatcgcag cagtgaggtg 1920 1980 agggetgeec gestestest segtegetge segetetggg gggatgesas ttgsctssag 2040 2100 ctggccatgc aagctgacgc ccgtgccttc tttgcccagg atggggtaca gtctctgctg acacagaagt ggtggggaga tatggccagc actacaccca tctgggccct ggttctcgcc 2160 ttettttgcc etccaetcat etacaecege etcateacet teaggaaate agaagaggag 2220 2280 cccacacggg aggagctaga gtttgacatg gatagtgtca ttaatgggga agggcctgtc 2340 gggacggcgg acccagccga gaagacgccg ctgggggtcc cgcgccagtc gggccgtccg 2400 ggttgctgcg ggggccgctg cggggggcgc cggtgcctac gccgctggtt ccacttctgg ggcgcgccgg tgaccatctt catgggcaac gtggtcagct acctgctgtt cctgctgctt 2460 ttetegeggg tgetgetegt ggattteeag eeggegeege eeggeteeet ggagetgetg 2520 2580 ctctatttct gggctttcac gctgctgtgc gaggaactgc gccagggcct gagcggaggc gggggcagcc tcgccagcgg gggccccggg cctggccatg cctcactgag ccagcgcctg 2640 egectetace tegeogacag etggaaceag tgegacetag tggeteteae etgetteete 2700

-46-

ctagacatag	gctgccggct	gaccccgggt	ttgtaccacc	tgggccgcac	tgtcctctgc	2/60
atcqacttca	tggttttcac	ggtgcggctg	cttcacatct	tcacggtcaa	caaacagctg	2820
gggcccaaga	tcgtcatcgt	gagcaagatg	atgaaggacg	tgttcttctt	cctcttcttc	2880
ctcqqcqtgt	ggctggtagc	ctatggcgtg	gccacggagg	ggctcctgag	gccacgggac	2940
agtgacttcc	caagtatcct	gcgccgcgtc	ttctaccgtc	cctacctgca	gatcttcggg	3000
cagattcccc	aggaggacat	ggacgtggcc	ctcatggagc	acagcaactg	ctcgtcggag	3060
cccggcttct	gggcacaccc	tcctggggcc	caggcgggca	cctgcgtctc	ccagtatgcc	3120
aactggctgg	tggtgctgct	cctcgtcatc	ttcctgctcg	tggccaacat	cctgctggtc	3180
aacttgctca	ttgccatgtt	cagttacaca	ttcggcaaag	tacagggcaa	cagcgatctc	3240
tactggaagg	cgcagcgtta	ccgcctcatc	cgggaattcc	actctcggcc	cgcgctggcc	3300
ccqcccttta	tcgtcatctc	ccacttgcgc	ctcctgctca	ggcaattgtg	caggcgaccc	3360
cggagccccc	agccgtcctc	cccggccctc	gagcatttcc	gggtttacct	ttctaaggaa	3420
qccqaqcqqa	agctgctaac	gtgggaatcg	gtgcataagg	agaactttct	gctggcacgc	3480
gctagggaca	agcgggagag	cgactccgag	cgtctgaagc	gcacgtccca	gaaggtggac	3540
ttggcactga	aacagctggg	acacatccgc	gagtacgaac	agcgcctgaa	agtgctggag	3600
cgggaggtcc	agcagtgtag	ccgcgtcctg	gggtgggtgg	ccgaggccct	gagccgctct	3660
gccttgctgc	ccccaggtgg	gccgccaccc	cctgacctgc	ctgggtccaa	agactgagcc	3720
ctactaacaa	acttcaagga	gaagccccca	caggggattt	tgctcctaga	gtaaggctca	3780
tctgggcctc	ggcccccgca	cctggtggcc	ttgtccttga	ggtgagcccc	atgtccatct	3840
gggccactgt	caggaccacc	tttgggagtg	tcatccttac	aaaccacagc	atgcccggct	3900
cctcccagaa	ccagtcccag	cctgggagga	tcaaggcctg	gatcccgggc	cgttatccat	3960
ctggaggctg	cagggtcctt	ggggtaacag	ggaccacaga	cccctcacca	ctcacagatt	4020
cctcacactg	gggaaataaa	gccatttcag	aggaaaaaaa	a		4061

<210> 30 <211> 1214 <212> PRT

<213> Homo Sapiens

<400> 30 Met Val Val Pro Glu Lys Glu Gln Ser Trp Ile Pro Lys Ile Phe Lys 10 5 Lys Lys Thr Cys Thr Thr Phe Ile Val Asp Ser Thr Asp Pro Gly Gly 25 20 Thr Leu Cys Gln Cys Gly Arg Pro Arg Thr Ala His Pro Ala Val Ala 35 40 45 Met Glu Asp Ala Phe Gly Ala Ala Val Val Thr Val Trp Asp Ser Asp 50 55 60 Ala His Thr Thr Glu Lys Pro Thr Asp Ala Tyr Gly Glu Leu Asp Phe 65 70 75 80 Thr Gly Ala Gly Arg Lys His Ser Asn Phe Leu Arg Leu Ser Asp Arg
85 90 95 90 85 Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe 100 105 110 Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro 125 . 120 Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg 140 135 Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr 150 155 Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala 165 170 175 Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly 180 185 190 185 190 180 Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro 195 200 205 Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro 210 220 215 Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His 230

-47-

-	_		Gly	245					250					255	
Ile	Ser	Gln	Gln 260	Lys	Thr	Gly	Val	Gly 265	Gly	Thr	Gly	Ile	Asp 270	Ile	Pro
		275	Leu				280					285			
	290		Thr			295					300				
305			Ala		310					315					320
	_		Gly	325					330					335	
			Pro 340					345					350		
_		355	Thr				360					365			
_	370		Glu			375					380				
385	_		Ser		390					395					400
			Asn	405					410					415	
_			Trp 420					425					430		
		435	Asp	_			440					445			
	450		Gly			455					460				
465			Pro		470					475					480
			Ala	485					490					495	
			Pro 500					505					510		
_		515	Ala				520					525			
	530		Gly			535					540				
545			Leu		550					555					560
			Leu	565					570					575	
			Glu 580					585					590		
_		595	Leu				600					605			
	610		Arg			615					620				
625					630					635					Arg 640
			Arg	645					650					655	
			660					665					670		Val
		675	Leu				680					685			
	690	Trp	Ala			695					700				
Thr 705	Arg	Leu	Ile	Thr	Phe 710	Arg	Lys	Ser	Glu	Glu 715	Glu	Pro	Thr	Arg	Glu 720

-48-

Glu Leu			725					730					735	
Gly Thr		740					745					750		
Ser Gly	Arg 755	Pro	Gly	Cys	Cys	Gly 760	Gly	Arg	Cys	Gly	Gly 765	Arg	Arg	Cys
Leu Arg		Trp	Phe	His	Phe 775	Trp	Gly	Ala	Pro	Val 780	Thr	Ile	Phe	Met
Gly Asn 785	Val	Val	Ser	Tyr 790	Leu	Leu	Phe	Leu	Leu 795	Leu	Phe	Ser	Arg	Val 800
Leu Leu	Val	Asp	Phe 805	Gln	Pro	Ala	Pro	Pro 810	Gly	Ser	Leu	Glu	Leu 815	Leu
Leu Tyr		820					825					830		
Leu Ser	Gly 835	Gly	Gly	Gly	Ser	Leu 840	Ala	Ser	Gly	Gly	Pro 845	Gly	Pro	Gly
His Ala 850					855					860				
Asn Gln 865	_			870					875					880
Cys Arg			885					890					895	
Ile Asp		900					905					910		
Asn Lys	915					920					925			
Asp Val					935					940				
Gly Val 945				950			-		955					960
Ser Ile			965					970					975	
Gln Ile		980					985					990		
Cys Ser	995					1000)				100	5		
Gly Thr	0				1015	5				1020	0			
Val Ile 1025				1030)				1039	5				1040
Ala Met			104	5		_	_	105	0				1055	5
Tyr Trp		1060)				106	5				107)	
Pro Ala	1075	5				1080)				1089	5		
Leu Arg	0				1095	5				110	0			
Ala Leu 1105				1110)				111!	5				1120
Leu Leu			112	5				113	0				113	5
Ala Arg		1140)				114	5				115	0	
Gln Lys	1155	5				1160)				116	5		
Glu Gln 117	0				1175	5				118	0			
Val Leu 1185	Gly	Trp	Val	Ala 1190		Ala	Leu	Ser	Arg 119		Ala	Leu	Leu	Pro 1200

-49-

Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp 1205 . 1210

> <210> 31 <211> 4646 <212> DNA <213> Homo Sapiens

<400> 31 tegacecacg egteegeea egegteegee caegegteeg eecaegegte egeeeaegeg 60 tecgeccacg cgtecgggt gaaagmramy cmygcktsms aaaaaccgte acttaggaaa 120 agatgtcctt tcgggcagcc aggctcagca tgaggaacag aaggaatgac actctggaca 180 gcacceggac cetgtactee agegegtete ggagcacaga ettgtettae agtgaaageg 240 300 acttggtgaa ttttattcaa gcaaatttta agaaacgaga atgtgtcttc tttaccaaag attccaaggc cacggagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag 360 gcacccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaattteeta 420 ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggaag tatatacgtc 480 tgtcctgcga cacggacgcg gaaatccttt acgagctgct gacccagcac tggcacctga 540 aaacacccaa cotggtcatt totgtgaccg ggggcgccaa gaacttogco otgaagccgc gcatgcgcaa gatottcago oggotcatot acatogogca gtocaaaggt gottggatto 600 660 tcacgggagg cacccattat ggcctgatga agtacatcgg ggaggtggtg agagataaca 720 ccatcagcag gagttcagag gagaatattg tggccattgg catagcagct tggggcatgg 780 840 tetecaaccg ggacaccete atcaggaatt gcgatgctga gggctatttt ttageccagt 900 accttatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaag ctccggaatc 960 agctagagaa gtatatetet gagegeacta tteaagatte caactatggt ggcaagatee 1020 ccattgtgtg ttttgcccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca 1080 tcaaaaataa aattoottgt gtggtggtgg aaggotoggg ccagatogot gatgtgatog 1140 ctagcctggt ggaggtggag gatgccctga catcttctgc cgtcaaggag aagctggtgc 1200 1260 getttttacc ccgcacggtg tcccggctgc ctgaggagga gactgagagt tggatcaaat ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaaaatg gaagaagctg 1320 gggatgaaat tgtgagcaat gccateteet acgetetata caaagcette agcaccagtg 1380 agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg 1440 acttagecaa tgatgagatt ttcaccaatg accgccgatg ggagtetget gacettcaag 1500 aagtcatgtt tacggctctc ataaaggaca gacccaagtt tgtccgcctc tttctggaga 1560 atggettgaa ectaeggaag ttteteacec atgatgteet caetgaacte tteteeaace 1620 acticageae gettgtgtae eggaatetge agategecaa gaatteetat aatgatgeee 1680 tectcaegtt tgtetggaaa etggttgega aetteegaag aggetteegg aaggaagaea 1740 gaaatggccg ggacgagatg gacatagaac tccacgacgt gtctcctatt actcggcacc 1800 ccctgcaagc tctcttcatc tgggccattc ttcagaataa gaaggaactc tccaaagtca 1860 tttgggagca gaccaggggc tgcactctgg cagccetggg agccagcaag cttctgaaga 1920 ctctggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaatg 1980 agtacgagac ccgggctgtt gagctgttca ctgagtgtta cagcagcgat gaagacttgg 2040 cagaacagct gctggtctat tcctgtgaag cttggggtgg aagcaactgt ctggagctgg 2100 cggtggaggc cacagaccag catttcatcg cccagcctgg ggtccagaat tttctttcta 2160 agcaatggta tggagagatt tecegagaca ccaagaactg gaagattate etgtgtetgt 2220 ttattatacc cttggtgggc tgtggctttg tatcatttag gaagaaacct gtcgacaagc 2280 acaagaaget gettiggtae tatgiggegi tetteacete eccettegig gietteteet 2340 ggaatgtggt cttctacatc gccttcctcc tgctgtttgc ctacgtgctg ctcatggatt 2400 tecatteggt gecacacece ecegagetgg tectgtacte getggtettt gteetettet 2460 gtgatgaagt gagacagtgg tacgtaaatg gggtgaatta ttttactgac ctgtggaatg 2520 tgatggacac getggggett ttttacttca tagcaggaat tgtatttegg etceactett 2580 ctaataaaag ctctttgtat tctggacgag tcattttctg tctggactac attattttca 2640 ctctaagatt gatccacatt tttactgtaa gcagaaactt aggacccaag attataatgc 2700 tgcagaggat gctgatcgat gtgttcttct tcctgttcct ctttgcggtg tggatggtgg 2760 cetttggcgt ggccaggcaa gggatcetta ggcagaatga gcagcgctgg aggtggatat 2820 teegtteggt catetacgag ccctacetgg ccatgttegg ccaggtgece agtgacgtgg 2880 atggtaccac gtatgacttt gcccactgca ccttcactgg gaatgagtcc aagccactgt 2940 gtgtggaget ggatgageac aacetgeeee ggtteeeega gtggateace ateceeetgg 3000 3060 tgtgcatcta catgttatcc accaacatcc tgctggtcaa cctgctggtc gccatgtttg

WO 00/40614 PCT/US99/29996

```
gctacacggt gggcaccgtc caggagaaca atgaccaggt ctggaagttc cagaggtact
                                                                      3120
                                                                      3180
tcctggtgca ggagtactgc agccgcctca atatcccctt ccccttcatc gtcttcgctt
acttctacat ggtggtgaag aagtgcttca agtgttgctg caaggagaaa aacatggagt
                                                                      3240
cttctqtctg ctgtttcaaa aatgaagaca atgagactct ggcatgggag ggtgtcatga
                                                                      3300
aggaaaacta ccttgtcaag atcaacacaa aagccaacga cacctcagag gaaatgaggc
                                                                      3360
                                                                      3420
atcgatttag acaactggat acaaagctta atgatctcaa gggtcttctg aaagagattg
ctaataaaat caaataaaac tgtatgaact ctaatggaga aaaatctaat tatagcaaga
                                                                      3480
tcatattaag gaatgctgat gaacaatttt gctatcgact actaaatgag agattttcag
                                                                      3540
acccctgggt acatggtgga tgattttaaa tcaccctagt gtgctgagac cttgagaata
                                                                      3660
aagtgtgtga ttggtttcat acttgaagac ggatataaag gaagaatatt tcctttatgt
                                                                      3720
qtttctccag aatggtgcct gtttctctct gtgtctcaat gcctgggact ggaggttgat
agtttaagtg tgttcttacc gcctcctttt tcctttaatc ttatttttga tgaacacata
                                                                      3780
tataggagaa catctatcct atgaataaga acctggtcat gctttactcc tgtattgtta
ttttgttcat ttccaattga ttctctactt ttcccttttt tgtattatgt gactaattag
                                                                      3900
                                                                      3960
ttggcatatt gtwaaaagtc tctcaaatta ggccagattc taaaacatgc tgcagcaaga
ggaccccgct ctcttcagga aaagtgtttt catttctcag gatgcttctt acctgtcaga
                                                                      4020
                                                                      4080
ggaggtgaca aggcagtete ttgetetett ggaeteacea ggeteetatt gaaggaacea
cocccattcc taaatatgtg aaaagtcgcc caaaatgcaa ccttgaaagg cactactgac.
                                                                      4140
tttgttctta ttggatactc ctcttattta ttattttcc attaaaaata atagctggct
                                                                      4200
attatagaaa atttagacca tacagagatg tagaaagaac ataaattgtc cccattacct
                                                                      4260
taaggtaatc actgctaaca atttctggat ggtttttcaa gtctattttt tttctatgta
                                                                      4320
tgtctcaatt ctctttcaaa attttacaga atgttatcat actacatata tacttttat
                                                                      4380
gtaagctttt tcacttagta ttttatcaaa tatgttttta ttatattcat agccttctta
                                                                      4440
aacattatat caataattgc ataataggca acctctagcg attaccataa ttttgctcat
                                                                      4500
                                                                      4560
tgaaggctat ctccagttga tcattgggat gagcatcttt gtgcatgaat cctattgctg
tatttgggaa aattttccaa ggttagattc caataaatat ctatttatta ttaaaaaaaa
                                                                      4620
                                                                      4646
aaaaaaagg gcggccgctc tagagt
```

<210> 32

<211> 1104

<212> PRT

<213> Homo Sapiens

<400> 32

Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg Arg Asn Asp 10 Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser Arg Ser Thr 25 20 Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile Gln Ala Asn 35 40 Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser Lys Ala Thr 55 Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly 70 Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys 90 8.5 Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly 105 100 Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile 125 120 115 Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu 140 135 Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro Arg 150 155 Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser Lys Gly 165 170 Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met Lys Tyr Ile 190 185 Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser Glu Glu Asn 200

-51-Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Cys Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Glu Gly Ser Gly Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala Gly . 390 Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg

-52-

Asp	Thr 690	Lys	Asn	Trp	Lys	Ile 695	Ile	Leu	Cys	Leu	Phe 700	Ile	Ile	Pro	Leu
705			Gly		710					715					720
			Leu	725					730					735	
			Trp 740					745					750		
Ala	Tyr	Val 755	Leu	Leu	Met	Asp	Phe 760	His	Ser	Val	Pro	His 765	Pro	Pro	Glu
	770		Tyr			775					780	_			
785	_	_	Val		790					795					800
			Leu	805					810					815	
			Ser 820					825					830		
Cys	Leu	Asp 835	Tyr	Ile	Ile	Phe	Thr 840	Leu	Arg	Leu	Ile	His 845	Ile	Phe	Thr
	850	-	Asn		-	855	_				860		_		
865			Phe		870					875					880
			Ala	885					890					895	
			Phe 900					905					910		
		915	Pro				920					925			
	930		Thr			935					940				
945			Leu		950					955					960
			Met	965					970					975	
			Gly 980					985					990		
	_	995	Phe				1000)				100	5		
	1010)	Pro			1015	5				1020	0			
		Lys	Cys	Phe			Cys	Cys	Lys			Asn	Met	Glu	
1025		C	C	nh -	1030		C1	7 0 0	7-2	103		T 013	ת 1 ת	Tro	104
			Суѕ	1045	5				1050)				105	5
			Lys 1060)				106	5				1070)	
		1079					1080	3				108	5		
Leu	Asn 1090		Leu	Lys	Gly	Leu 109		Lys	Glu	Ile	Ala 110	Asn O	Lys	Ile	Lys

In' itional Application No PCT/US 99/29996

A. CLASSII	FICATION OF SUBJECT MATTER C1201/68	8 C12N5/10 C07K	16/28
IPC 7	G01N33/53 A61K38/17	_ 5100, 15 55/10	
According to	o International Patent Classification (IPC) or to both national classific	ation and IPC	
B. FIELDS	SEARCHED		
Minimum do IPC 7	currentation searched (classification system followed by classification C12N C07K C12Q A61K G01N	ion symbols)	
	ion searched other than minimum documentation to the extent that s		
Electronic da	ata base consulted during the international search (name of data ba	ase and, where practical, search terms used)
BIOSIS ABS, C	, EPO-Internal, WPI Data, PAJ, MEDL HEM ABS Data, STRAND, GENSEQ, EMBL	INE, SCISEARCH, EMBASE,	BIOTECHNOLOGY
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the re-	elevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 16 February 1998 (1998-02-16) STRAUSBERG,R.: "ob70f05.s1 NCI_CO Homo sapiens cDNA clone IMAGE:13: mRNA sequence"	GAP_GCB1 36737 3',	1,2, 6-19, 25-35
	XP002138823 Accession AA809355		
х	DATABASE GENEMBL 'Online! 10 July 1998 (1998-07-10) MARRA ET AL.: "ub28d10.r1 Soares musculus cDNA clone IMAGE:137905! sequence" XP002149803 Accession AI050262	2NbMT Mus 9 5' mRNA	1,6-19, 25-35
		-/ 	
	·		
X Furti	her documents are listed in the continuation of box C.	X Patent family members are listed	in annex.
° Special ca	tegories of cited documents :	"T" later document published after the inte	mational filing date
"A" docume	ent defining the general state of the art which is not lered to be of particular relevance	or priority date and not in conflict with cited to understand the principle or th Invention	the application but
	document but published on or after the international	"X" document of particular relevance; the considered novel or cannot	t be considered to
L docume which	ent which may throw doubts on priority claim(s) or is cited to establish the publication date of another	involve an inventive step when the do	claimed invention
"O" docume	n or other special reason (as specified) ent referring to an oral disclosure, use, exhibition or	cannot be considered to involve an in document is combined with one or mo ments, such combination being obvio	ore other such docu-
other r *P* docume later tr	means ent published prior to the international filing date but aan the priority date dalmed	in the art. *a" document member of the same patent	
	actual completion of the international search	Date of mailing of the international sec	arch report
10	6 October 2000	3 0. 10	O. 00
Name and n	nailing address of the ISA	Authorized officer	
	European Patent Office, P.B. 5818 Patentiaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,	ALCONADA RODRIG	. A
l	Fax: (+31-70) 340-3016	ALCOHADA NODRIG	,

6

In' Itional Application No PCT/US 99/29996

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT				
tegory °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.		
{	DATABASE GENEMBL 'Online! 19 July 1997 (1997-07-19) STRAUSBERG, R.: "ni64ell.sl NCI_CGAP_Prl2 Homo sapiens cDNA clone IMAGE:981644 mRNA sequence" XP002148641 Accession AA523749	1,3, 10-19, 25-35		
(WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16)	1,4, 6-19, 25-35		
Y	page 4, line 7 -page 5, line 13 page 5, line 24 -page 7, line 28 SEQ ID NOs. 9 and 25	20-24		
x	WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27)	1,4, 6-19, 25-35		
	page 7, paragraph 2 page 9, paragraphs 2,3 page 13 -page 17 page 21, paragraph 3	20.04		
'	SEQ ID NOS: 109 and 112	20-24		
x	DATABASE GENEMBL 'Online! 18 November 1997 (1997-11-18) STRAUSBERG, R.: "nt76b07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204405, mRNA"	1,5-19, 25-35		
Y	XP002148642 Accession AA654650	20-24		
Y	DATABASE GENEMBL 'Online! 30 November 1998 (1998-11-30) SHIMIZU, N.: "Homo sapiens mRNA complete cds." XP002148643 Accession number AB001535 -& NAGAMINE ET AL.: "Molecular cloning of a novel putative Ca2+ channel protein (TRPC7) highly expressed in brain" GENOMICS, vol. 54, 15 November 1998 (1998-11-15), pages 124-131, XP000938744 the whole document	20-24		
	-/			

Int dional Application No PCT/US 99/29996

(Couring	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
ategory *	Citation of document, with Indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZHU, XI ET AL: "Molecular cloning of a widely expressed human homologue for the Drosophila trp gene." FEBS LETTERS, (1995) VOL. 373, NO. 3, PP. 193-198., XP000907241 page 194; figures 1,3	20,21, 23,25, 26,28, 29,31
1	HUNTER JOHN J ET AL: "Chromosomal localization and genomic characterization of the mouse melastatin gene (Mlsn1)." GENOMICS NOV. 15, 1998, vol. 54, no. 1, 15 November 1998 (1998-11-15), pages 116-123, XP000910696 ISSN: 0888-7543 cited in the application page 119; figure 2	20,21,23
4	WES PAUL D ET AL: "TRPC1, a human homolog of a Drosophila store-operated channel." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 1995, vol. 92, no. 21, 1995, pages 9652-9656, XP002138820 ISSN: 0027-8424 the whole document	20,21, 23,25, 26,28, 29,31
1	ZHU, XI ET AL: "Trp, A novel mammalian gene family essential for agonist-activated capacitative Ca-2+ entry." CELL, vol. 85, no. 5, 1996, pages 661-671, XP000907242 page 662 page 665 figures 1,5,6	20,21, 25,26, 28,29,31
	GARCIA REYNALDO L ET AL: "Differential expression of mammalian TRP homologues across tissues and cell lines." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1997, vol. 239, no. 1, 1997, pages 279-283, XP002138822 ISSN: 0006-291X See Materials and Methods figure 1	25,26, 28-30

In ational Application No
PCT/US 99/29996

		PCT/US 99/29996
C.(Continua	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SINKINS WILLIAM G ET AL: "Functional expression of TrpC1: A human homologue of the Drosophila Trp channel." BIOCHEMICAL JOURNAL APRIL, 1998, vol. 331, no. 1, April 1998 (1998-04), pages 331-339, XP000864583 ISSN: 0264-6021 page 333-335; figures 3-5	24
A	PREUSS KLAUS-DIETER ET AL: "Expression and characterization of a trp1 homolog from rat." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS NOV. 7, 1997, vol. 240, no. 1, 7 November 1997 (1997-11-07), pages 167-172, XP002138821 ISSN: 0006-291X figure 2	24
A	OBUKHOV, ALEXANDER G. ET AL: "Direct activation of trpl cation channels by G-alpha-11 subunits." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, (1996) VOL. 15, NO. 21, PP. 5833-5838., XP000907243 figure 2	24
P,X	WO 99 09199 A (RYAZANOV ALEXEY G ;PAVUR KAREN S (US); HAIT WILLIAM N (US); UNIV M) 25 February 1999 (1999-02-25) see melanome kinase polynucleotide and polypeptide sequences on page 16-17	1,3, 10-19, 25-36
P,X	WO 99 09166 A (SHAPERO MICHAEL H ;DENDREON CORP (US); LAUS REINER (US); TSAVALER) 25 February 1999 (1999-02-25) page 17, line 24 -page 18, line 9 page 25, line 19-32 page 28, line 1-4 SEQ ID NOS: 27, 28 and 31.	1,5-19, 25-35
Т	SCHARENBERG A M ET AL: "MLSN-1/SOC-1 defines a widely expressed Ca2+/cation channel family involved in Ca2+ homeostasis and store-operated Ca2+ signaling." FIFTY-THIRD ANNUAL MEETING OF THE SOCIETY OF GENERAL PHYSIOLOGISTS; WOODS HOLE, MASSACHUSETTS, USA; SEPTEMBER 9-11, 1999, vol. 114, no. 1, July 1999 (1999-07), page 14a XP000910708 Journal of General Physiology July, 1999 ISSN: 0022-1295	

..iemational application No.
PCT/US 99/29996

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2. X Claims Nos.:
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)
This International Searching Authority found multiple inventions in this international application, as follows:
see additional sheet As a result of the prior review under R. 40.2(e) PCT, no additional fees are to be refunded.
As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. X As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: 1-36
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Remark on Protest The additional search fees were accompanied by the applicant's protest. No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 1-5, 10-13, 16-19, 32-35 relate to an extremely large number of possible polynucleotides, polypeptides encoded by them, binding polypeptides, and kits and pharmaceutical compositions containing said polypeptides and polynucleotides. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the polynucleotide of SEQ ID NOs: 1, 27, 29 and 31 and the corresponding polypeptide of SEQ ID NOs: 2, 28, 30 and 32.

Present claims 16 and 17 relate to an extremely large number of possible compounds, namely, a polypeptide that binds to the polypeptide of the invention. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to an antibody, antibody fragment, F(ab)2 fragment or a fragment including a CDR3 region selective for the polypeptides of the invention.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1, 6-36 (partially) and 2 (complete)

An isolated nucleic acid molecule comprising a nucleic acid molecule that hybridizes to a nucleic acid molecule of SEQ ID NO:1 and which code for a SOC/CRAC polypeptide, nucleic acid molecules that differ in codon sequence due to degeneracy of the genetic code and complement thereof, polynucleotides which are not identical to the SEQ ID or sequences of GenBank accession number of Table 1; expression vector, host cells; polypeptide encoded thereof (SEQ ID NO:2); polypeptides binding to the polypeptide of SEQ ID NO:2, including antibodies; kits comprising agents that selectively bind to the polynucleotide (SEQ ID NO:1) or polypeptide (SEQ ID NO:2) of the invention; pharmaceutical compositions containing the polynucleotide or polypeptides of the invention; a method for isolating the SOC/CRAC molecule having SOC/CRAC calcium channel activity comprising contacting a binding molecule that is SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing SOC/CRAC molecules allowing the formation of the complex, detecting the formation of the complex, isolating the SOC/CRAC molecule and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity; a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity; a method to determine the level of SOC/CRAC expression in a subject, including expression of SOC/CRAC polypeptide or mRNA in a tissue or biological fluid sample using PCR, Northern blotting, and mono- and polyclonal antisera and a method for identifying agents useful in the modulation of the SOC/CRAC polypeptide kinase activity, comprising the use of aminoacids 999-1180 from SEQ ID NO:4 as a candidate kinase.

2. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:3 and to the encoded polypeptide of SEQ ID NO:4

3. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:5 and to the encoded polypeptide of SEQ ID NO:6

4. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:7 and to the encoded polypeptide of SEQ ID NO:8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1,6-36 (partially) and 37 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:23 and to the encoded polypeptide of SEQ ID NO:24 $\,$

6. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:25 and to the encoded polypeptide of SEQ ID NO:26 $\,$

7. Claims: 1,10-36 (partially) and 3 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:27 and to the encoded polypeptide of SEQ ID NO:28 $\,$

8. Claims: 1,6-36 (partially) and 4 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:29 and to the encoded polypeptide of SEQ ID NO:30 $\,$

9. Claims: 1,6-36 (partially) and 5 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:31 and to the encoded polypeptide of SEQ ID NO:32.

Information on patent family members

In dional Application No PCT/US 99/29996

Patent document cited in search report			Publication date		atent family member(s)	Publication date	
WO 9	9815657	A	16-04-1998	US	5919638 A	06-07-1999	
				EP	0954599 A	10-11-1999	
				US	6110675 A	29-08-2000	
WO 9	9837093	A	27-08-1998	AU	6181898 A	09-09-1998	
				CN	1252837 T	10-05-2000	
				EP	1005546 A	07-06-2000	
				NO	994069 A	22-10-1999	
				PL	335348 A	25-04-2000	
				ZA	9801585 A	04-09-1998	
WO 9	9909199	Α	25-02-1999	AU	9110098 A	08-03-1999	
WO 9	9909166		25-02-1999		9021898 A	08-03-1999	
		•••	•••••	EP	1005549 A	07-06-2000	

This Page Blank (uspto)